Title
What's the Catch?: Archaeological application of rapid collagen-based species identification for Pacific Salmon

Author names and affiliation

Kristine Korzow-Richter[1,2*], Krista McGrath[1*], Edouard Masson-MacLean[3], Simon Hickinbotham[1], Andrew Tedder[4], Kate Britton[3,5], Zoe Bottomley[1], Keith Dobney[3,6,7], Ardern Hulme-Beaman[7,8], Margherita Zona[3,7], Roman Fischer[9], Matthew J. Collins[10,11], Camilla F. Speller[1,6,12]

[1] BioArCh, Department of Archaeology, University of York, UK YO10 5DD
[2] Department of Archaeology, Max Planck Institute for the Science of Human History, Kahlaische Strasse 10, 07745 Jena, Germany
[3] Department of Archaeology, University of Aberdeen, St. Mary's Building, Elphinstone Road, Aberdeen, AB24 3UF, United Kingdom
[4] School of Chemistry & Bioscience, University of Bradford, Bradford, UK
[5] Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103, Leipzig, Germany
[6] Department of Archaeology, Simon Fraser University, Burnaby, B.C. V5A 1S6, Canada
[7] Department of Archaeology, Classics and Egyptology, University of Liverpool, L69 7WZ, UK
[8] Research Centre in Evolutionary Anthropology and Palaeoecology, School of Natural Sciences and Psychology, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK
[9] Target Discovery Institute, Nuffield Department of Medicine, University of Oxford
[10] University of Cambridge, McDonald Institute for Archaeological Research, Downing St, Cambridge CB2 3ER, UK
[11] Department of Biology, University of Copenhagen, Øster Farimagsgade 5, 1353 København K, Denmark
[12] Department of Anthropology, University of British Columbia, Canada

* These authors contributed equally to this work

Corresponding Authors: Edouard Masson-MacLean & Camilla Speller

Keywords
ZooMS, Collagen Peptide Mass Fingerprinting, Pacific Salmon, Yup'ik, Nunalleq

Abstract (approx 200 words)
Pacific salmon (*Oncorhynchus* spp.) are ecological and cultural keystone species along the Northwest Coast of North America and are ubiquitous in archaeological sites of the region. The inability to morphologically identify salmonid post-cranial remains to species, however, can limit our understanding of the ecological and cultural role different taxa played in the seasonal subsistence practises of Indigenous groups in the past. Here, we present a rapid, cost-effective ZooMS method to distinguish salmonid species based on collagen peptide mass-fingerprinting. Using modern reference material and an assemblage of 28 DNA-identified salmonid bones from the pre-contact Yup'ik site of Nunalleq, Western Alaska, we apply high-resolution mass spectrometry (LC-MS/MS) to identify a series of potential collagen peptide markers to distinguish Pacific salmon. We then confirm these peptide markers with a blind ZooMS analysis (MALDI-TOF-MS) of the archaeological remains. We successfully distinguish five species of anadromous salmon with this ZooMS approach, including one specimen that could not be identified through ancient DNA analysis. Our biomolecular identification of chum (43%), sockeye (21%), chinook (18%), pink (7%) and coho (11%), confirm the exploitation of all five available species of salmonid at Nunalleq.

1 Introduction

On the Northwest Coast of North America, Pacific salmon and trout (*Oncorhynchus* spp.) are both ecological (Willson and Halupka, 1995) and cultural (sensu Garibaldi and Turner, 2004) keystone species. Salmon become increasingly abundant in Northwest coast archaeological sites from 4000 BP, and likely played an important role in the emergence and development of distinctive cultures along the Northwest Coast. Multiple anadromous species of salmonids were exploited by Northwest Coast cultures, including the chinook or spring salmon (*Oncorhynchus tshawytscha*), sockeye (*Oncorhynchus nerka*), coho (*Oncorhynchus kisutch*), pink (*Oncorhynchus gorbuscha*), chum (*Oncorhynchus keta*), steelhead (*Oncorhynchus mykiss*) — as well as its non-anadromous form, rainbow trout — and cutthroat trout (*Oncorhynchus clarkii*). These salmonid species vary considerably in their distribution, life histories, spawning age and habitat, average size, fat content, etc. (Quinn, 2018) making species-level identification of salmon bone important for a wide range of anthropological and archaeological questions (S. Campbell and Butler, 2010), including fishing technologies, season of site occupation, sedentism and storage, distribution of resources within and between communities, and the rise of social complexity.

Pacific salmon and trout belong to the larger group of *Salmoniformes* including other highly fished species such as Atlantic salmon and trout (*Salmo spp.*) (Figure 1). *Salmoniformes* and *Escoiformes* diverged around 100 million years ago (MYA) (Broughton et al., 2013; Campbell et al., 2013; Santini et al., 2009; Zhivotovsky, 2015). Before the *Salmoniformes* diversified, a whole genome duplication occurred in the lineage (Allendorf and Thorgaard, 1984; Pasquier et al., 2016). Species within the order have experienced asymmetrical gene loss from this duplication and therefore still express both versions of many of the genes (Berthelot et al., 2014), with all individuals within any given species expressing either one or both versions of a gene. Thus, although *Salmoniformes* are well studied, their phylogenetic tree is not well resolved particularly near branches that likely diverged quickly or are polyphyletic (Crête-Lafrenière et al., 2012; Horreo, 2017; Lecaudey et al., 2018). Nevertheless, there is consensus that *Oncorhynchus* is a monophyletic genus that diverged from the sister taxa *Salvelinus* 18-30 MYA (Horreo, 2017; Lecaudey et al., 2018; Zhivotovsky, 2015). Within *Oncorhynchus* there are the masu salmon (*O. masou*) whose range covers the Western Pacific waters near China, Japan, and Korea, the Pacific trout (rainbow trout, steelhead, and cutthroat trout), and the anadromous Pacific salmon (chinook, sockeye, coho, pink, and chum) (Gong et al., 2017; Horreo, 2017; Kitano et al., 1997). These three groups diverged 10-25 MYA and the relationships within each group are well resolved (Alexandrou et al., 2013; Lecaudey et al., 2018; Rasmussen et al., 2009;

Zhivotovsky, 2015). There are a number of other non-anadromous *Oncorhynchus* trout species and subspecies in North America whose phylogenetic history is poorly resolved (Abadía-Cardoso et al., 2015; Penaluna et al., 2016; Petre and Bonar, 2017; Saglam et al., 2017). The non-anadromous species of Pacific trout are known to hybridize with each other as they can spawn at the same time and location (Allen et al., 2016; Ferguson et al., 1985; Kovach et al., 2018). The anadromous Pacific salmon, however, have a much lower hybridization rate particularly in the Pacific Northwest (Rasmussen et al., 2009; Rosenfield et al., 2000), because although they occupy the same waters as adults, spawning is highly specific to particular time periods and watersheds (Dittman and Quinn, 1996; Groot, 1991; Quinn and Dittman, 1990). This highly predictable spawning time and behaviour, linked to specific natal streams and watersheds, is what makes Pacific salmon species interesting from an archaeological perspective, in essence allowing for the seasonality of catch to be determined.

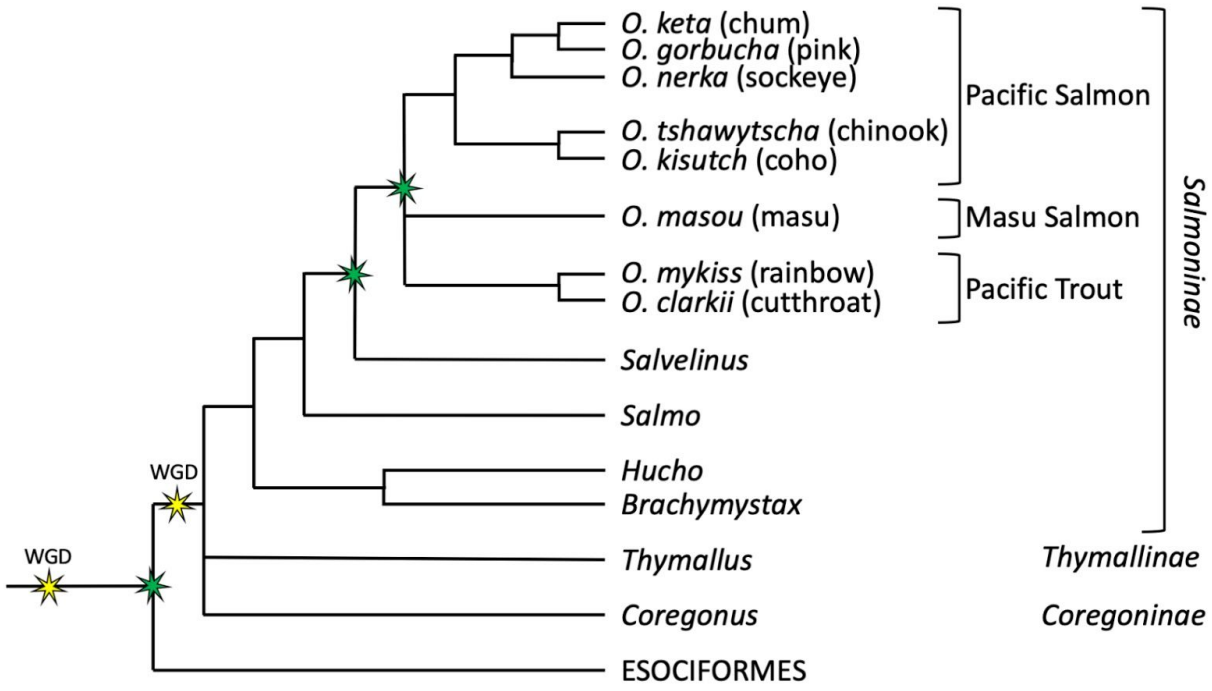

Figure 1: Consensus phylogenetic tree of *Salmoniformes* from (Horreo, 2017; Lecaudey et al., 2018; Zhivotovsky, 2015). The genus *Parahucho* is not included as its position is highly variable (Crête-Lafrenière et al., 2012) and not relevant for this study. Yellow stars indicate whole genome duplications (WGD) which occurred basal to all bony fish around 320 MYA (Gistelinch et al., 2016; Meyer and Van de Peer, 2005) and basal to *Salmoniformes* 70-90 MYA (Allendorf and Thorgaard, 1984; Pasquier et al., 2016). Green stars indicate key divergence dates for *Salmoniformes* and *Esociformes* at around 100 MYA, *Salvelinus* and *Oncorhynchus* 18-30 MYA, and the three groups within *Oncorhynchus* 10-25 MYA (Alexandrou et al., 2013; Horreo, 2017; Lecaudey et al., 2018).

Due to their close evolutionary relationship, however, distinguishing species of Pacific salmon using skeletal elements alone is extremely challenging (Butler and Bowers, 1998) and usually only possible if all elements are present, particularly those of the cranium (Cannon, 1988). Species discrimination has been attempted on vertebrae (Cannon, 1988), scales (Campbell, 2010; Henry, 1961; McMurrich, 1913), and otoliths (Casteel, 1974; Murray, 1994), by correlating incremental growth rings with known spawning ages of salmon. Radiographic analysis of vertebral growth rings proved a potentially non-destructive, rapid and cost-effective identification method (Cannon, 1988; Cannon and Carlson, 1991). Recently, however, Hofkamp and Butler (2017) demonstrated that vertebral rings identified through radiographic analysis do not represent annual growth rings, but internal structural walls uncorrelated with the known age of study fish. For scales and otoliths, the width and number of microscopic annuli more accurately correspond with annual cycles of fresh and salt-water occupation (Friedland and Haas, 1996; Fukuwaka, 1998; Murray, 1994; Saito et al., 2007). These tissues, however, are rarely recovered in representative numbers from zooarchaeological assemblages. Huber et al. (2011) also applied morphometric methods to vertebrae to address this identification problem; they measured vertebral length and height (and calculated the corresponding ratio) from modern Washington State reference specimens and classified them into discrete size/species classes. When applied to icthyoarchaeological remains from Alaska (and validated through ancient DNA analysis), however, this morphometric technique was only effective at discriminating pink salmon, and less so for other species (Moss et al., 2014).

Biomolecular approaches have been deployed in an effort to definitively identify archaeological salmonids —the most common being ancient DNA techniques (Cannon et al., 2011; Ewonus et al., 2011; Halffman et al., 2015; Kemp et al., 2014; Speller et al., 2005). These studies have typically targeted mitochondrial DNA (mtDNA), amplifying relatively short diagnostic fragments (150–300 bp) of the control region, cytochrome b (cytb) or 12S rRNA genes to identify taxa by comparison to a databank of known species and populations (Kemp et al., 2014; Yang and Speller, 2006). As previously mentioned, ancient DNA-based species identification methods have also been employed to validate zooarchaeological identification methods, for example osteometric/morphometric analyses of vertebral size and shape (Grier et al., 2013; Moss et al., 2014).

Recently, collagen-based identification approaches have emerged as a rapid, high-throughput and cost-effective alternative for taxonomic identification (Buckley, 2018; Collins et al., 2010). Collagen peptide mass fingerprinting (also known as ZooMS or Zooarchaeology by Mass Spectrometry) has been developed for the rapid

identification of archaeological bone (Buckley et al., 2009), and over the last 10 years has been used to discriminate morphologically similar ruminants (Taylor et al., 2018; von Holstein et al., 2014), marine mammals (Biard et al., 2017; Buckley et al., 2014; Hofman et al., 2018) and rodents (Buckley et al., 2016; Prendergast et al., 2017) from archaeological contexts. ZooMS uses collagen type I, a structural, triple helix fibral collagen that makes up the majority of the organic component of living bone. In tetrapods collagen type I is composed of two identical α1 chains (col1a1) and one α2 chain (col1a2). Collagen type I in many fish species is heterotrimeric and composed of three different chains. Col1a2 in fish is homologous to col1a2 in tetrapods. Col1a1 and col1a3 in fish are likely the result of the fish-specific whole genome duplication ~320 MYA and are both homologous to col1a1 in tetrapods (Christoffels et al., 2004; Gistelinck et al., 2016; Hurley et al., 2006; Kimura et al., 1991; Kimura and Ohno, 1987; Meyer and Van de Peer, 2005; Morvan-Dubois et al., 2003; Near et al., 2012). As there are three genes present which have higher mutation rates than mammal collagen (Martin and Palumbi, 1993; Takezaki, 2018), the potential for identifying species-specific collagen peptide markers among closely-related (and thus often morphological similar) fish taxa is quite high (Korzow-Richter et al., 2011). For example, a recent study demonstrated the utility of ZooMS for distinguishing anadromous Atlantic salmon from brown trout (Harvey et al., 2018). In order to explore the extent to which species-specific biomarkers (i.e., distinctive peptide peaks) could be identified in Pacific salmon —and to develop a rapid and cost-effective method for Pacific salmon identification— we applied ZooMS to an assemblage of salmonid remains excavated from the archaeological site of Nunalleq, Western Alaska.

1.1 Nunalleq Site Background

Nunalleq (GDN-248) is a pre-contact Yup'ik site, located on the shores of the Yukon-Kuskokwim (Y-K) Delta, close to the modern village of Quinhagak. Nunalleq was a densely occupied village, inhabited primarily during the 16th and 17th centuries AD (Ledger et al., 2018). Today, the region is home to the Central-Alaska Yupiit, who are likely descendants of the Thule —an archaeological culture found across the New World Arctic between ~1000 and 1400 AD (Raghavan et al., 2014). While historically under-researched, the sparse precontact Yup'ik archaeological record has recently been enriched by the collaborative, community led archaeological excavations at the Nunalleq site (Fienup-Riordan et al., 2015; Hillerdal, 2017; Hillerdal et al., 2018).

In partnership with the Native Corporation Quanirtuuq Inc., excavations at the site between 2009-2010 and 2012-2015 revealed multiple phases of occupation, with episodes of architectural remodelling (Ledger et al., 2018), prior to the eventual destruction of the site. The site includes the remains of a large sod and timber dwelling

comprising multiple rooms running off a central passageway (Figure 2). At least three different occupation phases, dating between c. 1570 and 1675 AD have been identified based on the presence of successive house floor and leveling deposits and evidence of architectural remodeling (Ledger et al., 2018, 2016). Phase II represents the last occupation at the site, dated approximately between 1640/1660 to 1675 AD, including dramatic evidence of destruction related to an attack on the village during a period of regional warfare (Fienup-Riordan and Rearden, 2016; Funk, 2010). Phase III was an earlier occupation, which began between AD 1620-1650 and endured for up to 35 years, during which the layout of the dwelling was very similar to Phase II (Ledger et al., 2018). Phase IV represents an earlier occupation that has only recently been excavated (Rick Knecht, *pers comm*) but for which initial dates suggest an occupation beginning cal 1570–1630 AD (Ledger et al. 2018). A combination of permafrost and waterlogging has led to the impressive preservation of organic materials at the site such as grass, leather, wood, bone and antler artefacts. Also preserved are a number of fishing related items such as barbed prongs, fish lures, net sinkers, nets, and wooden fish traps, suggesting that a variety of methods were used to exploit fish as observed during the historic period (Rick Knecht, *pers comm*). Human hair, animal fur, animal bone, including articulated salmon remains, and caribou antler were all recovered from the site, usually in very good condition, resulting in the preservation of collagen and ancient DNA thus providing reliable material for the application of biomolecular methods (Britton et al., 2018, 2013; Farrell et al., 2014; Gigleux et al., 2019; McManus-Fry et al., 2018; Raghavan et al., 2014). The faunal assemblage from the site was dominated by salmonids, most likely Pacific salmon and trout (*Oncorhynchus spp*.), marine mammals (mainly seals), caribou (*Rangifer tarandus*) and domestic dog (*Canis familiaris*), indicative of a tripartite resource subsistence strategy centered on salmonids, seals and caribou and for which dogs also played an important role for hunting and transportation as well as acting as a source of food in times of scarcity (Masson-MacLean et al., 2019; Masson-Maclean et al., n.d. In review; McManus-Fry et al., 2018). Isotope studies on human hair at the site also confirm the importance of salmonids and marine mammals in the diet (Britton et al., 2018, 2013).

An initial assessment of the hand-collected salmonid remains from the site highlighted that vertebrae represented 81% of the sample (N=2586). Very few non-salmonid remains have also been observed in bulk samples, currently undergoing analysis, suggesting other fish species were also exploited at the site, though very likely of lesser importance than salmonids. Isotope analyses also mirror this apparent predominance of salmonids over other fish resources, such as marine fish, in the diet (Britton et al., 2018). According to the ethnographic record, apart from salmonids, coastal historic Yupiit in SW Alaska traditionally exploited a variety of marine, anadromous and

freshwater fish from their environment throughout most of the year, such as salmonids, smelt, herring, blackfish, pike or flounder (Barker, 1993; Frank, 2008; LaVine et al., 2007; Wolfe et al., 1984). Salmonids present in the region include Pacific salmon and trout (*Oncorhynchus* spp.) —but excluding cutthroat trout which does not occur today north of the Alaska Peninsula (Hauser, 2014)— char (*Salvelinus spp.*), whitefish and cisco (*Coregoninae spp.*), and Arctic grayling (*Thymallus arcticus*). At the time of writing, the Nunalleq fish assemblage appears to be almost exclusively composed of salmonid remains with Pacific salmon the only species identified to date (Masson-MacLean et al., 2019).

The current study applied ancient DNA analyses to: 1) validate the species of Pacific salmon present in the Nunalleq archaeological assemblage; and 2) use these data to identify potential diagnostic collagen peptide markers for five Pacific salmon species.

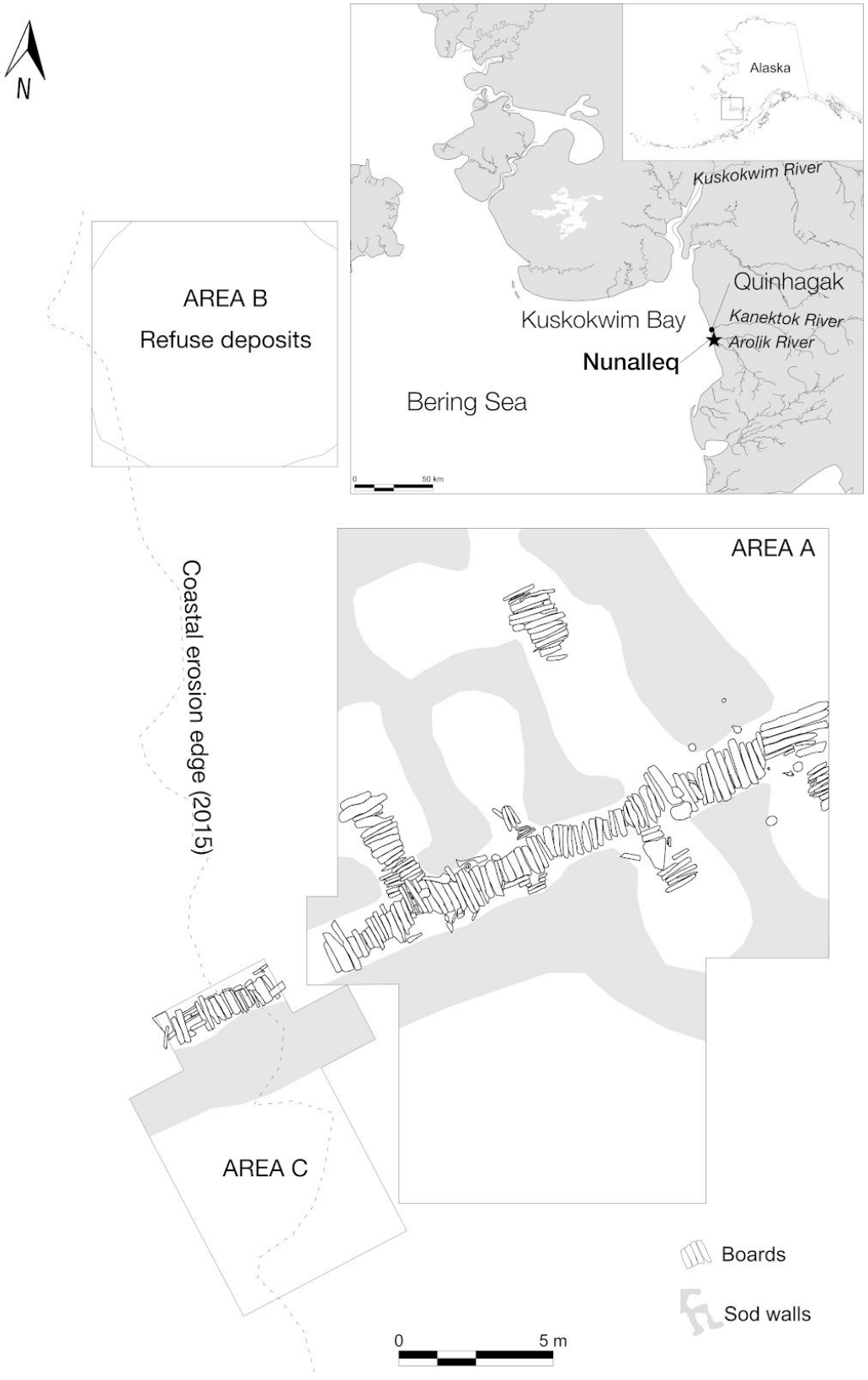

Figure 2: Site plan (Phase II layout) and location map of the Nunalleq site.

2. Materials and Methods

2.1 Archaeological Specimens

A sample of 28 salmon caudal vertebrae were analysed in this study originating from different house floor deposits in order to reduce the possibility of selecting elements from the same individual (Table 1). To maximise the chance of all species of Pacific salmon being represented in the sample, vertebrae of various sizes were selected. The vertebrae were all hand-collected during the excavation and the fish remains, along with the rest of the Nunalleq faunal assemblage, are currently stored at the University of Aberdeen, Scotland.

Table 1: Sample proveniences and taxonomic identifications based on mtDNA and peptide markers.

Identifications of the Archaeological Samples with aDNA and ZooMS

Lot #	Sample ID	Cytb ID	D-loop ID	ZooMS ID	Final Identification
15869	14083.91.F2	Chum	Chum	Chum	Chum
15870	14049.53.F1	Sockeye		Sockeye	Sockeye
15871	13141.59.F3	Sockeye	Sockeye	Sockeye	Sockeye
15872	13141.59.F2	*Fail*	*Chum*	*Chum*	*Chum*
15873	13159.56.F1	Chum		Chum	Chum
15874	12009.59.F1	Pink	Pink	Pink	Pink
15875	13170.57.F1	Chum		Chum	Chum
15876	12014.29.F1	Chinook	Chinook	Chinook	Chinook
15877	13159.56.F3	Chum		Chum	Chum
15878	13189.56.F1	Chum	Chum	Chum	Chum
15879	13141.59.F1	Sockeye		Sockeye	Sockeye
15880	13092.20.F1	Chum		Chum	Chum
15881	13159.56.F2	Chum		Chum	Chum
15882*	14049.53.F2	Sockeye		Sockeye	Sockeye
15883	14083.91.F4	Chum		Chum	Chum
15884	13139.223.F1	Coho	Coho	Coho	Coho
15885	13130.227.F1	Chinook		Chinook	Chinook
15886	14083.91.F3	Sockeye		Sockeye	Sockeye
15887*	13188.57.F1	Chum		Chum	Chum
15888	13189.56.F2	Coho	Coho	Coho	Coho
15889	12014.41.F1	Chinook	Chinook	Chinook	Chinook
15890	14105.25.F1	*Fail*		*Chum*	*Chum*
15891	14083.91.F1	Chinook		Chinook	Chinook
15892	12014.41.F2	Sockeye		Sockeye	Sockeye
15893*	13093.37.F1	Chinook		Chinook	Chinook
15894*	13189.56.F3	Coho	Coho	Coho	Coho
15895*	14069.67.F1	Pink	Pink	Pink	Pink
15896	13047.58.F1	Chum	Chum	Chum	Chum

*Samples were used for ZooMS biomarker discovery and LC-MS/MS.

Sample ID is the unique identifier assigned to the assemblage at the University of Aberdden.

Lot Number is the unqiue identifer assigned by the BioArCh laboratory at the University of York.

Reference specimens

We obtained modern reference specimens of rainbow trout as a representative of the Pacific trout group. We also obtained modern Atlantic salmon and brown trout as representatives of a different genus in order to compare the collagen sequences and peptide markers of Pacific salmon with other salmonids. The Atlantic salmon and brown trout were acquired from the "Proyecto Arca" (Arca Project) carried out by the Asociación de pescadores *Las Mestas del Narcea* (an association of fishermen in Pravia, Asturia, North Spain) where salmon were legally caught, transported to a hatchery where the eggs were gathered during spawning, and then released back to the river. When salmon died naturally during the project, the fish were frozen and sent to Madrid to Dr. Laura Llorente Rodríguez. The skeletons were prepared using water maceration. We obtained modern reference specimens of rainbow trout from the Kanektok River near Quinhagak, Alaska; the trout was collected in the summer of 2015 by members of Qanirtuuq Incorporated, and forms part of the zooarchaeological reference collection at the University of Aberdeen.

2.3 Genetic Species Identifications

2.3.1 DNA Extraction, Amplification, and Sequencing

The 28 archaeological salmon bones from the Nunalleq site initially underwent ancient DNA analysis to confirm species identity. Sample preparation and DNA extraction were conducted in the BioArCh dedicated Ancient DNA laboratory at the University of York, following strict contamination control protocols such as: the separation of the pre-PCR and post-PCR workspaces; the use of ancient DNA dedicated equipment including clothing, equipment and reagents; the analysis of both blank extractions and negative controls alongside the ancient DNA samples; and multiple extractions from the same sample in order to reproduce the original results. First, all bone samples were split in half, with one subsample retained for ZooMS and the other for DNA analysis. The subsamples for DNA analyses were chemically decontaminated through submersion in 6% sodium hypochlorite for 5 minutes followed by UV irradiation for 30 min on two sides, before being crushed into powder. The samples were incubated overnight in 1 mL of lysis buffer (0.5 M EDTA pH 8.0 and 0.5 mg/mL proteinase K) in a rotating hybridization oven at 50°C. Samples were then centrifuged and the supernatant from each sample was concentrated to <100 µL using Amicon Ultra-4 Centrifugal Filter Devices (10 KD, 4mL, Millipore). Concentrated extracts were purified using QIAquick spin columns (QIAGEN, Hilden, Germany) based on the method developed by Yang et al. (1998); 60 µL of DNA from each sample was eluted from the QIAquick column for PCR amplification.

PCR amplifications targeted two fragments of mtDNA (cytb and D-loop) previously demonstrated to accurately distinguish Pacific salmon species (Yang and Speller, 2006). PCR amplifications were performed in a 30 µL reaction volume containing 1.5X Applied Biosystems™ Buffer, 2 mM $MgCl_2$, 0.2 mM dNTP, 1.0 mg/mL BSA, 0.3 mM primer, 2.0-3.0 µL DNA sample and 2.5-3.75 U AmpliTaq Gold LD (Applied Biosystems). Initially, samples were amplified with primers targeting 168 bp of the cytb gene; a subset of samples were also amplified with a longer fragment targeting 249 bp of the hypervariable D-loop to confirm cytb species identifications. Successfully amplified samples were sequenced using forward and/or reverse primers at Eurofins, Germany. Good quality sequences were obtained for all but two of the amplified samples: 13141.59.F2 and 14105.25.F1 yielded poor quality cytb sequences, however, the D-loop sequence for 13141.59.F2 was sufficient to make a species identification. Thirty-eight sequences were uploaded to the Genetic Sequence Database at the National Center for Biotechnical Information (NCBI) (GenBank ID: MK863994-MK864031).

2.3.2 Species Identification

The obtained sequences were first compared to Genbank reference sequences through the online BLAST application to determine their closest match, and to ensure that they did not match with any other unexpected species or sequences. Sample sequences were visually edited and base pair ambiguities were examined using ChromasPro software (www.technelysium.com.au). Alignments of the ancient DNA sequences and published salmonid reference sequences were conducted using ClustalW (Thompson et al., 1994), through BioEdit (Hall, 1999). Species identifications were confirmed through phylogenetic analysis of a 118 bp cytb fragment and 198 bp D-loop fragment, respectively. The ModelTest (ver. 2.3) software (Posada and Crandall, 1998) was employed to determine the best-fit model (GTR+G, selected by AIC), implemented in MrBayes 3.2.5 (Ronquist and Huelsenbeck, 2003). Ten million generations of analyses were performed to produce the phylogeny and clade credibility scores, with a burnin of one million generations, chosen *post hoc* after examination of parameter convergence in Tracer 1.7 (Rambaut et al., 2018). Phylogenetic trees were created using FigTree 1.4.4 (Rambaut, 2007).

2.4 Collagen Sequencing and Peptide Mass Fingerprinting (ZooMS)

2.4.1 Experimental design and workflow

Following genetic species identification, one representative (i.e., one archaeological bone subsample) from each of the five DNA-identified species was selected for collagen sequencing (subsequently referred to as ‘archaeological reference samples’) (Table 1).

From the modern samples, one Atlantic salmon, two brown trout, and one rainbow trout were selected randomly for sequencing (subsequently referred to as ‘modern reference samples’). First, collagen from the archaeological and modern reference samples was extracted, enzymatically digested, and assessed for overall quality and preservation via matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS). The resulting MALDI spectra along with publicly accessible protein sequence data were used to generate a list of ‘theoretical peptide markers’ (detailed methods below). Second, collagen from the reference samples was analysed via nanoflow liquid chromatography-tandem mass spectrometry (nLC-MS/MS) to authenticate the sequences from the theoretical peptide markers and generate the list of shared and diagnostic peptide markers. Third, collagen from all remaining archaeological samples, two Atlantic salmon, two brown trout, and one rainbow trout was extracted and characterized using MALDI-TOF-MS. In a blind test, mass spectra from the five newly extracted modern samples and all 28 archaeological samples (including the five archaeological reference samples) were compared against the list of diagnostic peptide markers to identify the species. Finally, in a blind test, species identifications based on the peptide markers were compared to those obtained using genetic methods; Figure 3 displays the complete workflow for biomarker discovery and confirmation.

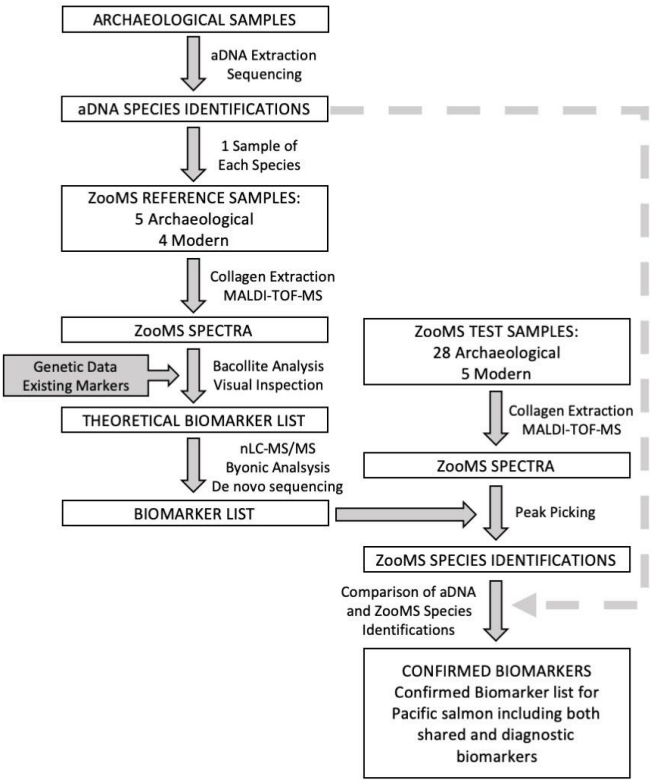

Figure 3:Workflow diagram illustrating sequential steps of marker discovery and confirmation.

2.4.2 Collagen Extraction

For the modern reference samples, between 50-200 mg of salmon bone was placed in eppendorf tubes and demineralized in 8 mL of 0.6 M hydrochloric acid at 4°C. Following centrifugation, the supernatant was removed and the samples were rinsed three times in 8 mL of ultrapure water. Then, 4 mL of HCl (pH 3.0) was added to the samples, followed by incubation for 48 hours at 80°C to gelatinise the collagen. The collagen was then purified using an ultrafilter (Amicon Ultra-4 Centrifugal Filter Devices, 30 KD, 4 ml, Millipore) and then freeze dried. Between 0.1-0.4 mg of dried collagen was resuspended in 50 µL of 50 mM ammonium bicarbonate solution (NH_4HCO_3) pH 8.0 (AmBic) and incubated overnight at 37°C with 0.4 µg of trypsin. A second resuspension of collagen was incubated overnight at 25°C with 0.4 µg of elastase.

For the archaeological reference samples, between 10-30 mg of bone was placed in eppendorf tubes and demineralized in 250 µL of 0.6 M hydrochloric acid at 4°C. Following centrifugation, the supernatant was removed and 200 µl of 0.1 M NaOH was added to the sample to remove humics and other chromophoric compounds. The samples were vortexed and centrifuged, and the supernatant discarded. The samples were rinsed three times in 200 µL of AmBic. Finally, 100 µL of AmBic was added to the samples, followed by incubation for one hour at 65°C to gelatinise the collagen. Samples were centrifuged and 50 µL of the supernatant was incubated overnight at 37°C with 0.4 µg of trypsin (Buckley et al., 2009). An additional 50 µL of gelatinized collagen from each sample was incubated overnight at 25°C with 0.4 µg of elastase.

Following trypsin/elastase digestion, both modern and archaeological samples were then acidified to 0.1% trifluoroacetic acid (TFA), and purified using 100 µL C18 resin ZipTip® pipette tips (EMD Millipore) using conditioning and eluting solutions composed of 50% acetonitrile and 0.1% TFA, and 0.1% TFA for the lower hydrophobicity buffer. Collagen was eluted in 50 µL.

2.4.3 MALDI-TOF-MS and theoretical biomarker determination

The archaeological and modern reference samples were analyzed via MALDI-TOF-MS in order to create a list of theoretical peptide markers, and subsequently by LC-MS/MS to confirm those biomarkers. For MALDI-TOF-MS, 1 µL of the trypsin-digested extract was mixed with 1 µL of α-cyano-hydroxycinnamic acid and spotted in triplicate with calibration standards onto a 384 spot MALDI target plate. The samples were run on a Bruker ultraflex III MALDI TOF/TOF mass spectrometer with a Nd:YAG smart beam laser, with a SNAP averaging algorithm was used to obtain monoiosotopic masses (C

4.9384, N 1.3577, O 1.4773, S 0.0417, H 7.7583). Spectra were visually inspected using the mMass software (Strohalm et al., 2008) and all samples produced high-quality spectra with high signal to noise ratios, and multiple discrete peaks.

Publically available salmon collagen sequence data was obtained from Genbank and Phylofish transcriptome database (Pasquier et al., 2016), including collagen type I for Atlantic salmon (XP_014048044.1, XP_014059932.1, XP_014033985.1, XP_013998297.1, XP_014035319.1, XP_014065361.1), brown trout (NSTid:64109-54964-73256, NSTid:2881, NSTid:28977, NSTid:31221-73256, NSTid:44848, NSTid:51479-73256-64109), rainbow trout (NP_001117649.1, CDQ73303.1, BAB55663.1, BAB55662.1, D2OMid:55140-44007, D2OMid:56516-54949-6383-55140), coho (XM_020493253.1, XP_020360620.1, XP_020355639.1, XP_020316038.1, XP_020327438.1), and chum (BAB79230.1, BAB79229.1). The sequences were aligned and compared to collagen type I sequences from the model organism zebrafish (*Danio rerio)* (AAH63249.1, AAH71278.1, AAH58045.1), in order to confirm which sequences were col1a1 and col1a3 as these are commonly mislabeled by annotation software (SI Table 5). The versions were then determined by similarity and randomly numbered. The Phylofish data was checked for completeness and both chimeric and possibly chimeric sequences were removed.

Using Bacollite v0.0.2 (Hickinbotham and Collins, 2017; Hickinbotham et al., n.d.), the sequence data was theoretically (i.e., *in-silico*) digested with trypsin and the resulting peptides were used to generate a set of theoretical spectral peaks which allowed for hydroxylation of proline and deamidation of glutamine as well as isotope distribution alignment (SI Methods 1). The theoretical dataset was then compared against the spectra data obtained from each of the species to detect 'theoretical peptide markers'. This list of theoretical peptide markers was augmented with published biomarkers for salmonids (Harvey et al., 2018) and peptide markers identified through visual comparison of spectra from the five archaeological reference samples.

2.4.4 nLC-MS/MS and confirmation of biomarkers

In order to confirm the sequences of the theoretical peptide markers, the archaeological and modern reference samples were characterized by nLC-MS/MS. The trypsin and elastase digests were combined in equal peptide concentrations for each of the modern and archaeological reference samples after digest, respectively, and analyzed at the Discovery Proteomics Facility, Target Discovery Institute, Oxford. High resolution mass spectrometry was performed on a Q-Exactive instrument after UPLC separation on an EASY-Spray column (50 cm, 75 mm ID, PepMap RSLC C18, 2 mm) connected to a Dionex Ultimate 3000 nUPLC (all Thermo Scientific) using a gradient of 2–35%

Acetonitrile in 0.1% Formic Acid/5% DMSO and a flow rate of 250 nL/min @40°C. MS spectra were acquired at a resolution of 70000 @200 m/z using an ion target of 3E6 between 380 and 1800 m/z. MS/MS spectra of up to f15 precursor masses at a signal threshold of 1E5 counts and a dynamic exclusion for 7 seconds were acquired at a resolution of 17500 using an ion target of 1E5 and a maximal injection time of 128 ms. Precursor masses were isolated with an isolation window of 1.6 Th and fragmented with 28% normalized collision energy. MS/MS data was uploaded to ProteomeXchange through MassIVE MSV000083687 (doi:10.25345/C5G04C).

For confirmation of biomarkers where sequence data was available, the MS/MS spectral data was then analyzed using Byonic v.3.2.0 (Protein Metrics Inc. (Bern et al., 2012)) with the following parameters: cleavage sites fully specific C-term R and K; 2 missed cleavages allowed; mass changes - 4 common, 1 rare; common: oxidation on M, P, and K, deamidation on N and Q; rare: pyro-Glu on N-term E and Q; no sequence variations allowed, wildcard search disabled. The CollagenDB database was formed from all of the sequence data for collagen type I, the top proteins in the search against the salmonid proteome (SI Table 2), and common contaminants (Harris et al., 2002; Keller et al., 2008). Validation tests were completed to identify the number of PTMs and the sequences used (described in SI Methods 2 and 3, SI Table 1, 2, and 3). The peptide markers which were shared between all species from the analysis of the MALDI data were confirmed through analysis of the Atlantic salmon and coho samples. The masses corresponding to the unique peptide markers were checked in all samples to find sequences.

Following this analysis, diagnostic peptide markers which still had no confirmed sequence data were analyzed through an error tolerant search strategy for identification of novel sequence variants using assisted *de novo* sequencing. This approach was validated using coho MS/MS data with rainbow trout sequences as the database in Byonic (described in SI Methods 4, SI Table 4, SI Figures 1 and 2) using the following less restrictive parameters to allow for identification of potential sequence variants: cleavage sites fully specific C-term R and K; 1 missed cleavage allowed; mass changes - 2 common, 2 rare; common: oxidation on K, M, P, deamidation on N, Q; rare: all sequence variants, wildcard search disabled. Following validation the MS/MS spectral data for three species (pink, sockeye, chum) were then reanalyzed with the same parameters, but with the database including all salmon type I sequences. To reduce the number of potential oxidations that Byonic needed to identify, we identified fixed proline oxidation sites through the results of the initial biomarker confirmation searches. We then fixed these sites as oxidized prolines in Byonic and reran the error tolerant

searches, thus increasing the possibility of finding mutations on peptides with a high number of proline oxidations.

Once all the putative biomarkers had sequence data, these locations were mapped back onto the full protein sequences for collagen type I using Geneious v.11.1.5 (https://www.geneious.com). The locations of the peptide markers were compared to known mammal peptide markers (A-G, P1, P2) (Buckley et al., 2014, 2009). The locations were also compared between the two versions of each gene to identify the masses of all sequence variants for all of the locations where peptide markers were present (SI Figure 5). The peptide marker sequences, as well as all sequence variants at those locations, were used to create a biomarker database for Byonic which consisted of: 1) the putative biomarker sequences and all of the sequence variants found in either the genetic or the error tolerant protein sequencing at those locations; 2) all of the collagen III and V genes that are publicly available for salmonids; 3) any protein that was in any of the top ten proteins found in the whole proteome validation; and 4) common contaminants. The MS/MS data from every species was then analyzed using Byonic (with the same parameters as the non-error tolerant Byonic run, but with 5 common, 1 rare mutations allowed). The proteins were filtered to a protein false discovery rate (FDR) of <1%. The peptides were filtered so that only peptides with a PEP 2D score less than 0.001 were considered (all of the peptides with a PEP 2D score had peptide FDR scores of less than 0.1%). A peptide marker was considered present in a species if the marker had a coverage depth of at least 3x with masses corresponding to the mass present in the MALDI spectra. These combined analyses resulted in a list of sequences for the shared and diagnostic peptide markers for the Pacific salmon species. The above steps were able to resolve the amino acid sequence for all but one of the biomarkers; only a single peptide marker at m/z 1936 (specific to coho) could not be confirmed. Additional analyses undertaken to resolve the sequence for this marker, including additional LC-MS/MS, MALDI-TOF-MS/MS and expanded database searches, are described in the SI Methods 6.

2.5 Biomarker Test

The diagnostic peptide markers identified above were validated in a blind test on the 28 archaeological samples. Collagen was extracted from the archaeological samples using the methods described above, with enzymatic digestion taking place with trypsin only. As above, 1 µL of extract was mixed with 1 µL of α-cyano-hydroxycinnamic acid and spotted in triplicate with calibration standards, and run on a Bruker ultraflex III MALDI TOF/TOF mass spectrometer. As there were no remains of the rainbow/steelhead trout identified in the Nunalleq archaeological remains, we used available sequence data and MALDI-TOF spectra from modern reference samples to confirm that the diagnostic

peptide markers could also differentiate the Pacific salmon from rainbow trout. Spectra were visually inspected using the mMass software (Strohalm et al., 2008) and poor quality spectra (i.e., with low signal to noise ratios, or few to no discrete peaks) were eliminated from the dataset. Spectra were then provided with new sample codes (by CFS) and compared against the list of prospective peptide markers in a blind test. Following independent identification by two different researchers (ZB, KKR), shared and diagnostic peptide markers which appeared in all samples of the same species were incorporated into Table 2, and the resulting ZooMS taxonomic identifications compared to the genetic identifications. MALDI data was uploaded to Zenodo RN 2649336 (doi:10.5281/zenodo.2649336).

3. Results

3.1 Ancient DNA Species Identifications

Mitochondrial DNA was successfully amplified and sequenced from 27 of the 28 archaeological samples, a success rate of 96%. Table 1 provides the mtDNA-based species identification results confirmed through phylogenetic analysis of the cytb and D-loop fragments (SI Figure 4). The contamination control procedures undertaken in this study were successful at eliminating any systematic contamination as no PCR amplification was observed in blank extracts or PCR negative controls. D-loop and cytb sequences from all samples yielded the same species identities, suggesting that there was no cross-contamination between samples.

3.2 Collagen Peptide Markers

3.2.1 Protein Sequence Data and Expression

Salmonids retain two different versions of each collagen type I gene from the whole genome duplication and some species are expressing both versions of the genes in their transcriptomes (Pasquier et al., 2016). Therefore, each species will have a total of six potential collagen genes rather than the usual three for fish and two for mammals and birds. We were able to reconstruct the entire expressed protein sequence for all six of the genes for Atlantic salmon, brown trout, and rainbow trout; five out of six sequences for coho and two for chum (SI Table 5). LC-MS/MS data indicated that both versions of all three genes are being expressed in every individual both for archaeological and modern samples (SI Methods 5, SI Figure 3). Therefore, all six proteins where available were used as the sequence input for Bacollite and Byonic.

3.2.2 Theoretical Biomarkers

Processing the MALDI-TOF spectra with Bacollite yielded around 300 combinations for theoretical peptide markers and candidate sequences, in addition to the 16 published

peptide markers (Harvey et al., 2018) and ten potential peptide markers identified through visual comparison of spectra from the five archaeological samples (SI Figure 6).

3.2.3 Shared Biomarkers

Comparison of the LC-MS/MS data from Atlantic salmon and coho with published sequence data confirmed 24 peptide markers from the MALDI-TOF spectra shared across the six Pacific salmon species, rainbow trout, brown trout, and Atlantic salmon tested in this study (Table 2). These markers were used to confirm both the overall quality of the spectra, and an identification to the salmonid family, before the diagnostic markers were used to distinguish the species. The majority of the peptide markers (n=18) have only a single peptide sequence that composes the peak at that m/z ratio. However, four of the peptide peaks (m/z 1359, 1520, 1536, 1600) have two sequences with the same mass that are present in the LC-MS/MS data. Peaks associated with all four of these masses were present in all of the MALDI-TOF spectra which is in agreement with the data from mammals that some of the primary peaks are actually composed of more than one peptide with similar masses. These masses should be used with caution for identification. An additional three masses (m/z 1185, 1800, 2472) that were present in all of the samples were not included in the list of 24 peptide markers because the peaks were composed of three or more different peptides, and the composition of the particular peptides differed between species. These masses should be included for future developmental work on salmonids using LC-MS/MS analysis, but because they are a complex mixture, they should not be used when only MALDI-TOF data is available.

Table 2: Shared peptide markers that are present in all species, and with efficient ionisation in the MALDI-TOF MS.

Shared Biomarkers Between Pacific Salmon Species and *Salmo salar*

Mass (m/z)	Sequence	# Oxidations	Peptide code	seqpos
882.5	LGPAGASGPR	0	COL1A2T69(2)*	895-904
912.5	GIVGLPGQR	1	COL1A1T69	958-966
1040.6	GGIGSVGPSGPR	0	COL1A3T47	685-696
1200.6	GQPGNIGFPGPK	2	COL1A2T36	485-496
1203.6	GQPGVMGFPGPK	2	COL1A1T38	574-585
1359.7[§]	GESGSFGPAGPAGLR	0	COL1A2T54	692-706
1359.7[§]	GFPGLAGQLGEPGK	2	COL1A1T71	970-983
1443.7[+]	GAAGPPGATGFPGAAGR	2	COL1A1T62*	865-881
1503.8	GDAGPSGLTGFPGAAGR	1	COL1A2T61*	776-792
1514.8	GNSGPAGSAGSQGAIGAR	0	COL1A2T45*	590-607
1520.8[§]/1536.8[§]	DGMSGLPGPIGPPGPR	1/2	COL1A1T87	1152-1167
1520.8[§]/1536.8[§]	GDKGLPGGPGAVGEPGR	0/1	COL1A2T68/69(1)*	878-894
1563.8	DGMNGLPGPIGPPGPR	2	COL1A3T87	1152-1157
1584.8/1600.8[§+]	GAPGSSGIAGAPGFPGPR	2/3	COL1A1T21	397-414
1596.8	GAAGLPGVAGAPGFPGPR	3	COL1A2T20	308-325
1600.8[§+]	GLGGDPGPSGQKGDSGAK	1	COL1A2T22	341-358
1764.8[+]/1780.8	GSPGPMGPPGLSGAPGEAGR	1/2	COL1A3T73	994-1013
1774.0[+]/1790.8	GPPGPMGPPGLAGAPGESGR	1/2	COL1A1T73*	994-1013
1878.9/1895.9	GNDGPMGAPGTPGPGGIAGQR	1/2	COL1A3T68	937-957
1934.0	SGDRGESGPAGPAGIAGPAGPR	0	COL1A1T78	1062-1083
2138.0	GEGGPAGPPGFAGPPGSDGQSGPR	2	COL1A2T56*	713-736
2167.0	GEAGAAGAPGGQGPPGLQGMPGER	3	COL1A1T49	709-732
2612.1/2628.1	GFTGMQGPSGPAGPSGESGPAGASGPAGPR	1/2	COL1A1T85*	1111-1140
2973.3/2989.3	GPPGPPGSSGPQGFTGPPGEPGEAGSSGPMGPR	1/2	COL1A1T3	187-219

*Markers previously identified in Harvey *et. al.* (2018).

[§]Masses where two peptides compose a peak. Masses where more than two peptides compose a peak have been removed.

[+]Masses where there is also a common contamiant that could be present.

Peptide code corresonds to the trypsin cutting locations along the Bovine collagen sequence from Buckley *et. al.* (2009). The sequence location corresponds to our alignment to BOVIN collagen P02453 (col1a1/col1a3) and P02465 (col1a2) (SI).

Five peptide marker peptides share masses (m/z 1443, 1600, 1764, 1774) with peptides found in the common laboratory contaminants for ZooMS: keratin and BSA. Three of these five have other levels of proline hydroxylation that are present in the MALDI-TOF spectra, but two only are present in areas with contaminants. The obtained LC-MS/MS data displayed very little contamination as well as no evidence for these potential contaminant peptides in our dataset; nevertheless, when using only MALDI-TOF spectra for species identification, contaminants and authentic collagen peptides cannot be easily differentiated. Therefore, these peptides are not recommended for sole use as ‘diagnostic’ species determination peptides to indicate a salmonid species, but instead used in conjunction with other markers to make taxonomic identifications.

Harvey et al. (2018) identified nine biomarkers shared between the two *Salmo* species. We find six of these present in all of the salmonid species we tested (COL1A1T62, COL1A2T45, COL1A2T61, COL1A1T73, COL1A2T56, COL1A1T85). An additional two markers were not visible in the MALDI-TOF spectra, but were present in the LC-MS/MS (COL1A1T47 and COL1A2T21). Therefore, differences in collagen extraction or preservation may account for the lack of these markers in the MALDI-TOF spectra, rather than true sequence differences. The final marker, COL1A2T69 is one of the nine

major diagnostic mammal markers (peptide D). This region of the collagen sequence contains two diagnostic mammal markers separated by a three amino acid peptide (SI Figure 5). In several families of fish including salmonids there is an arginine instead of a proline in the middle of COL1A2T69, resulting in two smaller markers. All of our samples contain the second half of COL1A2T69 from Harvey et al. (2018) at m/z 882. However, some fish including salmonids also have a lysine instead of an arginine as the amino acid before COL1A2T69. This appears to not be cleaved effectively in all sample preparation methods as Harvey et al. (2018) observe a peak at m/z 1220 which corresponds to the first half of COL1A2T69, while we find the peak at m/z 1520/1536 corresponding to a missed cleavage at the lysine before COL1A2T69 adding three amino acids of COL1A2T68 to the first half COL1A2T69.

Harvey et al. (2018) identified two diagnostic biomarkers for Atlantic salmon and brown trout. One of these markers, COL1A2T3, was not present in any of the samples, but as our samples generally had a poor resolution above m/z 3000 we would encourage further exploration of this mass area where possible for diagnostic markers. The other marker COL1A3T47 was present in our samples; all of the Pacific species we tested shared the sequence variant identified in brown trout at m/z 996.

3.2.4 Diagnostic Biomarkers

We identified three potential candidate markers for distinguishing between the salmonid genera *Salmo* and *Oncorhynchus* (COL1A2T56, COL1A2T53/54, and COL1A2T67). Although distinguishing between these two genera may not be relevant in many archaeological contexts as the ranges of these two genera rarely overlap, it may be a useful method for investigating historic samples and introduced species. We caution, however, that these potential markers should be confirmed through the analysis of additional *Salmo* species and specimens.

We identified fourteen markers that uniquely distinguish the six Pacific salmon species present in our modern and archaeological assemblages. We confirmed the sequence locations and variations using genetic sequence data for ten of the markers, we identified three additional markers through error tolerant *de novo* sequencing, and one sequence remains unresolved (Table 3, SI Figure 7). The marker with an unknown sequence (m/z 1936) is only present in coho. We present here two sequences from a theoretical digestion of collagen which could correspond to this marker as they show variation between the species in genetic data; nevertheless, they were unable to be confirmed through LC-MS/MS or MALDI-TOF-MS/MS. Furthermore, coho only has genetic data available for one version of the col1a1 and therefore, 1936 could correspond to an unknown variant of col1a1 which was unable to be identified through

error tolerant searching. We propose that the marker at 1936 derives from salmonid collagen type I as: 1) it is diagnostic in the archaeological samples, which displayed little evidence of contamination in the form of common contaminant sequences in the LC-MS/MS analysis; 2) it does not correspond to any other protein in the salmonid proteomes which were tested; and 3) the majority of the peaks in the MALDI-TOF spectra could be correlated with strict trypsin digestion of peptide fragments from collagen type I (and not other proteins) based upon the LC-MS/MS data. The clear detection of this peptide in the MALDI-TOF spectral data but not the LC-MS/MS data may reflect differences in ionization between MALDI and ESI (Stapels and Barofsky, 2004). Nevertheless, until the sequence for this species-specific marker can be confirmed (e.g., through genomic sequencing of coho collagen), we recommend that this marker be used with caution, and/or confirmed with other biomolecular methods (e.g., mtDNA or morphometric analysis).

Table 3: Diagnostic markers for the Pacific Salmonid species (bold) along with the other sequences at those positions with inefficient ionisation in the MALDI-TOF MS (§) or which may be non-diagnostic as multiple sequences with that mass fly in the MALDI-TOF MS (+).

Mass (m/z)	Sequence	# Oxidations	Peptide Code	seqpos	*O. tshawytscha* (chinook)	*O. nerka* (sockeye)	*O. kisutch* (coho)	*O. gorbuscha* (pink)	*O. keta* (chum)	*S. salar* (Atlantic salmon)	*O. mykiss* (rainbow trout)
996.5	GGIGS**A**GP**A**GPR*	0	COL1A3T47	685-696	X	X	X	X	X		X
1026.5	GGIGS**A**GP**T**GPR*	0								X	
1040.6	GGIGS**V**GP**S**GPR	0			X	X	X	X	X	X	X
1874.8	GP**A**GNIGMPGMTGTQGEAGR	1	COL1A2T70	905-924	X		X	X	X	X	X
1890.8	GP**S**GNIGMPGMTGTQGEAGR	1			X	X	X	X	X	X	
1936.0	*DGARGGPGPSGPPGPSGANGEK*	*0*	*COL1A2T52/53*	*670-691*			X				
	PGNRGESGPTGNGGPVGAAGAR	*1*	*COL1A2T75/76*	*879-918*							
2142.0	GE**S**GA**A**GALGG**P**GAPGLQGMPGER	3	COL1A3T49	709-732				X			
2160.0*	GE**S**GA**A**GA**P**GG**M**GAPGLQGMPGER	3			X	X	X		X	X	X
2174.0§	GE**T**GA**A**GA**P**GG**M**GAPGLQGMPGER	3								X	
2202.0§	GE**T**GV**A**GA**P**GG**M**GAPGLQGMPGER	3			X	X	X	X	X		X
2122.0	GEGGPAGPPGFAGPPGSDGQ**S**GPR*	1°	COL1A2T56	713-736		X			X	X	
2138.0	GEGGPAGPPGFAGPPGSDGQ**S**GPR*	2			X	X	X	X	X	X	X
2148.0	GEGGPAGPPGFAGPPGSDGQ**P**GPR	2								X	X
2152.0	GEGGPAGPPGFAGPPGSDGQ**T**GPR	2						X			
2456.2	GDVGPAGPAGP**A**GQSGPSGASGPAGPPGGR	1	COL1A2T60	746-775	X		X			X	X
2472.2*	GDVGPAGPAGP**S**GQSGPSGASGPAGPPGGR	1			X	X	X	X	X		X
2612.2/2628.2	GFTGMQGP**S**GP**A**GP**S**GESGPAG**A**SGPAGPR*	1/2	COL1A1T85	1111-1140	X	X	X	X	X	X	X
2669.2	GFTGMQGP**P**GP**S**GQ**S**GESGPAG**A**SGPAGPR	1			X		X	X	X	X	X
2699.2	GFTGMQGP**P**GP**S**GQ**S**GESGPAG**T**SGPAGPR	1				X					
2881.3	G**AA**GPSGPPGPSGANGEKGESGSFGPAGPAGLR	2	COL1A2T53/54	674-706	X	X	X	X	X		
2893.3§/**2909.3**	G**GP**GPSGPPGPSGANGEKGESGSFGPAGPAGLR	2/3			X						X
2922.4	G**PA**GPSGPPGPSGANGEKGESGSFGPAGPAGLR	3								X	X
2931.5§/**2947.5**	GP**T**GESGPPG**A**PGTAGPQGVLG**PS**GFVGLPGSR	1/2	COL1A2T67 Mammal Marker: G	845-877	X		X	X	X		X
2933.4/2949.4§	GP**S**GESGPPG**A**PGTAGPQGVLG**PS**GFVGLPGSR	2/3			X	X	X	X	X		X
2955.5*/**2971.5**	GP**T**GESGPPG**A**PGTAGPQGVLG**HA**GFVGLPGSR	1/2								X	
2967.5§/2983.5§	GP**S**GESGPPG**P**PGTAGPQGVLG**HA**GFVGLPGSR	1/2								X	

*Markers previously identified in Harvey *et. al.* (2018).

§Marker flies poorly in the MALDI-TOF MS, visible in LC-MS/MS only.

*Multiple sequences fly at this position and therefore this is not a marker.

°This is a potential species specific oxidation state and should be used with caution for identification.

Markers which were sequenced through error tolerant assisted *de novo* sequencing are indicated by bold Xs.

Shared markers which correspond to the same area of the sequence as diagnostic markers are indicated by light grey Xs.

Diagnostic MALDI markers are indicated by bold masses.

Sequences in light grey were not confirmed with LC-MS/MS and are candidate sequences only.

The markers highlighted in grey provide a unique species ID within Pacific Salmon.

Peptide code corresonds to the trypsin cutting locations along the Bovine collagen sequence from Buckley et. al. (2009). The sequence location corresponds to our alignment to BOVIN collagen P02453 (col1a1/col1a3) and P02465 (col1a2) (SI).

We were able to confirm the sequences of the other thirteen diagnostic peptide markers. Ten of the markers are clearly sequence variations. One of the markers (m/z

2122) corresponds to a taxa-specific oxidation state of the shared sequence between all species for COL1A2T56 (m/z 2138), which is present in sockeye, chum and Atlantic salmon. In combination with other markers, 2122 is important as a diagnostic marker for chum. In our dataset, this oxidation state was fixed across all individuals within sockeye, chum and Atlantic salmon. Nevertheless, it will be important to test the extent to which this is consistent in salmonids across a range of environments and time periods. Thus, we caution that the identification for chum should be first made using the shared salmonid markers, then by the diagnostic markers shared between chum and coho and the absence of the diagnostic coho marker (m/z 1936), finally the presence of a peak at 2122 should be considered. Ideally, a subset of the putative ZooMS-identified chum samples should have taxonomic identifications confirmed through ancient DNA analysis.

As salmonids express two versions of the collagen I genes, there are several combinations of the six genes that may lead to diagnostic biomarkers. First, both versions are identical in most species, but there is a variant in one version leading to a diagnostic peptide marker. This is the case for COL1A1T56 where all species have the same version at 2138, but pink salmon have a difference in one gene copy leading to an additional peak at 2152 (Figure 4a). Second, the versions are different from each other but shared between all the species, and one version has a variant leading to a diagnostic peptide marker. This is the case for COL1A1T85 where all species share one version with a mass at 2612. Most species share the second version at 2669, except for sockeye which have a distinctive marker at 2699 (Figure 4b). Third, there are variants of both versions which express in different ways in different species groups. This is the least common of the options, and it is mostly seen when including the genus *Salmo* in the analysis. The most obvious example of this case is COL1A3T49 where there are two versions, each of which has a variant that is present in one species. However, only one of the four variants is easily identifiable in the MALDI-TOF spectra (m/z 2142 present only in pink salmon) —the others were identified only in the LC-MS/MS data.

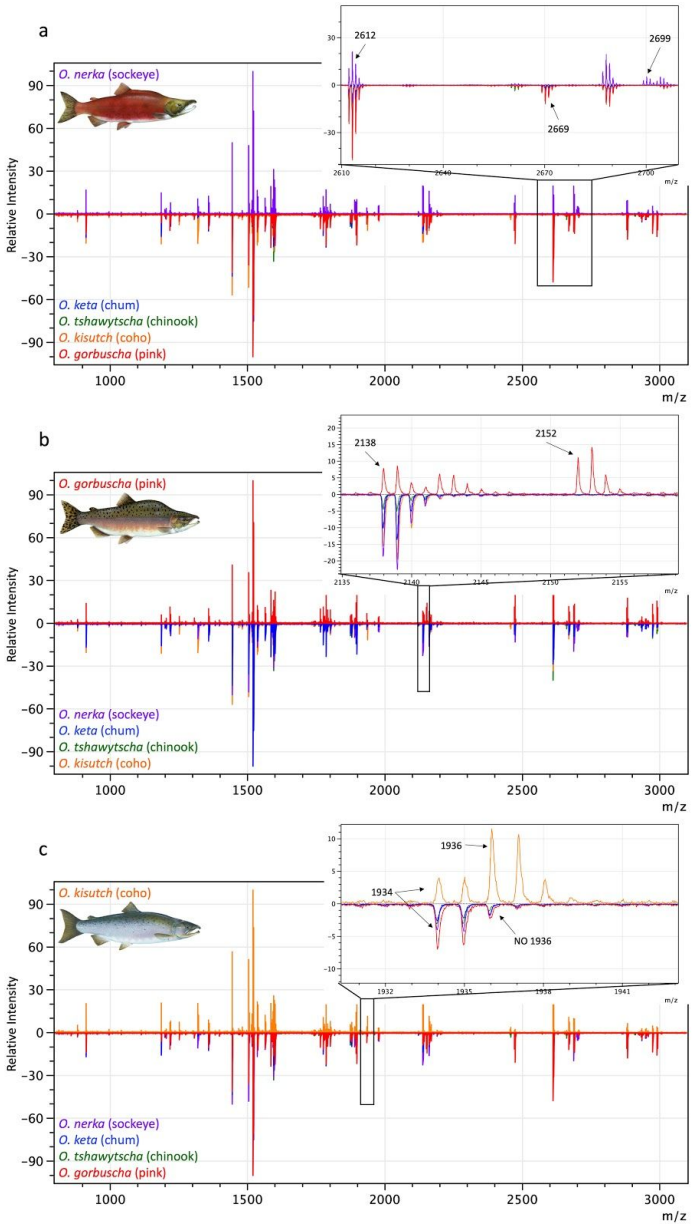

Figure 4: Examples of the diagnostic peptide markers in MALDI-TOF spectra. a) Close-up of peptide marker at peptide code COL1A1T85 for the Pacific Salmonid species. All five species share the one version of the peptide that flies at m/z 2612. Four of the species (lower spectra) share their second version which flies at m/z 2669. Sockeye (upper spectra) has a difference in that version which creates a species diagnostic peptide marker at m/z 2699. b) Close-up of peptide marker at peptide code COL1A2T56 for the Pacific Salmonid species. All five species share the peptide that flies at m/z 2138. Four of these species (lower spectra) have two copies of this version, with a substitution in one version in pink salmon (upper spectra) which yields a diagnostic peptide marker at m/z 2152. c) Close-up of peptide marker at m/z 1936 for the Pacific Salmonid species. All five species share the peptide COL1A1T78 that flies at m/z 1934. Coho (upper spectra) has a diagnostic peptide that flies at m/z 1936. The peak from this peptide overlaps the isotope distribution from the shared peak at m/z 1934, but clearly shows an increase in signal corresponding a new peptide and not the falling isotope distribution for the peptide with the primary peak at m/z 1934.

Given the relative similarity of collagen sequences between salmon species in comparison to other fish species, and the fact they are expressing two versions of all three collagen genes, unique biomarkers were not found for all species; nevertheless, unique *combinations* of markers were found for all species including predicted markers based upon sequence data for rainbow trout. These diagnostic peptide markers include a marker at 1936 found in coho. All species share a peptide marker at 1934 and the isotope peaks overlap the diagnostic marker at 1936, so care should be taken to assess the difference between a new peptide with a primary peak at 1936 and the secondary isotope peaks from the peptide with the primary peak at 1934 (Figure 4c).

3.3 Species Identification using Diagnostic Biomarkers

In a blind test, we assessed the reliability and replicability of identifying the suite of 24 shared peptide markers, and 14 diagnostic peptide markers in 28 archaeological salmonid bones. All of the shared markers were present in all 28 of the archaeological spectra, although some were present in quite low intensities (S/N ratio of 3 or relative intensity at 1%). Using the 14 diagnostic peptide markers, the blind test was successful in assigning species-level identification to all 28 of the archaeological specimens. For the diagnostic peptide markers, two of the archaeological salmonid samples (13141.59.F2 and 13130.227.F1) were each missing one diagnostic marker corresponding to their species, but successful species-level identification of all samples was still possible based on the combination of other peaks. Blind analysis of the archaeological salmon spectra by two different observers identified identical lists of shared and diagnostic peptide markers for each of the 28 archaeological samples, respectively. For the 27 archaeological samples identified to species through genetic analysis, ZooMS and DNA-based identifications produced the same result.

4 Discussion

4.1 Pacific Salmonid Biomarkers

We demonstrated that each salmon individual expresses both versions of their three collagen genes in bone (SI Methods 5). Further, we were able to identify patterns of diagnostic peptide markers that distinguish the Pacific Salmonid species. For five of the species (chinook, sockeye, coho, pink, and chum) these markers allowed successful identification of the archaeological specimens from Nunalleq. We noted excellent collagen preservation in the salmon remains from Nunalleq, with all of the shared and diagnostic peaks present in the majority of the remains. Thus, we achieved a high success rate for species identifications using the ZooMS approach. At archaeological

sites where collagen is not as well preserved, taxonomic identifications may not always be possible to the species level if not all diagnostic markers are present. We used MALDI-TOF spectra from a modern reference sample of rainbow trout to confirm that these diagnostic peptide markers could also differentiate rainbow/steelhead trout from other Pacific salmon, although we recommend that these diagnostic peptide markers should be validated on a larger assemblage of modern and/or archaeological rainbow trout specimens. Finally, we were able to identify several candidate markers for distinguishing between the genera *Salmo* and *Oncorhynchus*. As the number of *Salmo* species we tested was low, these candidate markers should be further investigated to assess whether these markers vary within the *Salmo* genus. Likewise, future research is necessary to examine the effectiveness of the ZooMS approach for distinguishing masu salmon in the Western Pacific, the Pacific trout, and the other *Oncorhynchus* non-anadromous species such as Gila (*O. gilae*) and Apache trout (*O. apache*).

4.2 Identification of collagen biomarkers for fish

ZooMS relies on the identification of taxa specific biomarkers. In mammals, the sequence data for ZooMS peptide markers have frequently been identified using a "bottom-up" approach. This approach generally uses MS/MS data to search proteomics databases (UniProt and SwissProt) for homologous proteins and assisted *de novo* sequencing to find single amino acid changes among closely related species (Buckley et al., 2009). When sequences are highly conserved, such as in the collagen of large mammals, minor variations in the collagen sequence are easily identified using this approach (Welker et al., 2015a, 2015b). Other groups, such as small mammals and fish, are more challenging because they have fewer published sequences available (Song and Wangs, 2013) (SI Figure 8) and faster mutation rates —thus more potential variation in collagen sequences (Gu and Li, 1992; Martin and Palumbi, 1993; Takezaki, 2018). This challenge is augmented for salmonids because they express six collagen genes (rather than three) due to their whole genome duplication. The "bottom-up" approach is more challenging when sequences are highly divergent, reference sequences from closely related species are unavailable, and/or the post-translational modifications (particularly proline hydroxylation) are less understood. As the number of potential amino acid differences and post-translational modifications increase, the computational time and power needed to analyze the data expands exponentially and the certainty for any individual amino acid substitution decreases. As several of the diagnostic peptide markers we found had three or four amino acid differences from each other, using the "bottom-up" approach, would be difficult and time intensive.

We, therefore, used a “sequence-guided” approach, starting with acquiring as much full sequence data for closely related species as possible including mining transcriptome databases for collagen sequences. As several salmonid species have complete genome sequences (or are highly studied due to their commercial value), we acquired collagen sequences for five salmonid species that allowed us to understand the variation within the collagen genes and between versions of genes. Then we used this data to identify a list of potential peaks present in the MALDI-TOF spectra and using this list we identified peptide markers that were unique to each species. This more targeted approach reduced the computational time for peptide marker identification. Subsequently, through visual comparison of the spectra, we identified five additional masses as potential markers in species that lacked published sequence data, and thus we only needed to attempt assisted *de novo* sequencing to identify a limited number of peptides with known mass values (m/z 1936, 2122, 2142, 2152, 2699). We confirmed sequence differences for three of these markers with the parameters tested (m/z 2142, 2152, 2699). In addition, we confirmed that one of the markers was a taxa-specific variation in oxidation state rather than a sequence difference (m/z 2122) and should therefore be used with caution during identification. When looking at a large number of diverse species (like different families of fish) this “sequence guided” approach both provides more robust results and requires less computational time. Nevertheless, the fact that we were unable to identify the sequence for peptide 1936 highlights the challenges in identifying diagnostic biomarkers for relatively understudied taxa like fish. As fish collagen databases continue to expand, and as full genome sequence data becomes available for more species, we will likewise expand the capacity to rapidly identify and confirm species-specific markers.

4.3 Comparison of new fish markers to published mammal markers

There are nine major markers that are used to identify mammal species using ZooMS (A - G, P1, P2) (Buckley et al., 2014, 2009). Two of the markers are on col1a1 in mammals. We checked these areas in both col1a1 and col1a3 in salmon. Marker F does show sequence differences, but none of these peptides are reliably present in the MALDI-TOF spectra. Harvey et. al (2018) identified marker P1 as present in the spectra from both col1a1 and col1a3. In col1a1 both versions are identical in all of the salmon species tested (COL1A1T47, m/z 1010). In col1a3 there are two versions in salmon. One is present in all species tested at m/z 1040. The other version is one of the diagnostic biomarkers found by Harvey et. al (2018). All of the Pacific salmon share the same variant with brown trout at m/z 996.

The other six markers for mammals are on col1a2. Two of these markers are found as either shared (C) or diagnostic (G) markers in salmon. Marker A is present in our

salmon samples, but its peak is composed of at least four different peptides, some of which have different expression between different species of salmon, and is therefore not suitable as a salmonid marker. Three of the markers (B and D) are affected by amino acid differences which interfere with trypsin digestion (SI Figure 5) and either create different markers or do not appear in the MALDI-TOF spectra. Given that one-third of the standard mammal markers are affected by mutations that affect trypsin digestion and all three collagen chains have roughly equally high mutation rates as opposed to the slower and differential mutation rates between col1a1 and col1a2 in mammals (Buckley, 2018; Buckley et al., 2008), a new set of standard markers need to be developed specifically for fish. In addition, given the diversity of fish sequences, we expect this to be a larger standard set than is necessary for mammals. Until a new standard set of markers is proposed and validated, the presence of at least half of the shared salmonid markers should be confirmed before species specific identification is attempted in order to ensure that the MALDI-TOF spectra in fact derives from salmonid collagen type I (SI Table 6).

4.4 Comparison of biomolecular approaches for taxonomic identification

In this study, taxonomic identifications obtained through genetic approaches were in complete agreement with those obtained through ZooMS. The success rate for both identification methods was extremely high: 27 of 28 archaeological bones were identified through mtDNA analyses (96%), while 100% of the samples were identified through ZooMS. Several studies have noted a high success rate for fish DNA amplification from archaeological contexts along the Northwest Coast (Cannon and Yang, 2006; Ewonus et al., 2011; Moss et al., 2016/8; Speller et al., 2012), suggesting that conditions are generally favorable for ancient DNA preservation in such Holocene deposits. Nevertheless, PCR inhibitors have been noted to significantly impact the success of some ancient DNA studies in this region (Grier et al., 2013; Kemp et al., 2014; Monroe et al., 2013; Palmer et al., 2018); in these contexts, ZooMS may be a more efficient species identification method than traditional DNA methods. Compared to ancient DNA analysis, ZooMS also offers additional advantages in terms of sample throughput and laboratory cost. As such, it is an ideal screening method for species-level identifications and estimations of overall biomolecular preservation ahead of high-resolution genomic studies of phylogeography (Brown et al., 2016; Star et al., 2017) or phenotype (Royle et al., 2018; Thompson et al., 2019).

ZooMS is also valuable because it can be used on fragmentary bone or from bone elements that can not be morphologically identified (such as rib bones). In many cases

these bones cannot be identified even to the family level or identifications differ between experts (Gobalet, 2001; LeFebvre and Sharpe, 2018). Here, we provide shared markers that could be useful for identifying bones first as Pacific Salmon, and subsequently to the species level. As ZooMS markers for fish are currently very limited, this approach is still in its infancy. However, once more markers are published, the resolution will drastically increase allowing fast and relatively inexpensive identification of fish bones when archaeological questions require a higher taxonomic resolution than morphological analysis can provide.

The ability to identify fish bones using a cost-effective method has implications for both archaeological and palaeoecological research. In archaeological contexts, this method could enhance our understanding of human economic strategies and resource exploitation. Species identification would also then facilitate other types of analyses, for example, stable isotope analysis for informing more nuanced isotopic mixing models of human palaeodiet. Such analyses could also inform palaeoecological research, not only allowing insight into salmon population biodiversity through time but also —when integrated with other analyses such as stable isotope research— the exploration of niche plasticity amongst different species. We show the great potential for ZooMS to work well on fish bones —a potential currently limited by the lack of reference MALDI-TOF spectra and collagen sequence data for taxa of zooarchaeological interest as well as the limited number of laboratories providing ZooMS identifications.

4.5 Implications for the Nunalleq Site

The inability to securely identify pacific salmon vertebrae to species using traditional zooarchaeological comparative methods restricts our understanding of fishing strategies and salmon exploitation at Nunalleq, a major component of subsistence not only in the archaeological past but also historically and into the present day in the Y-K Delta. This study has shown the promising potential of rapid collagen-based species identification to resolve a significant issue in archaeological research at Nunalleq, where salmon remains, particularly vertebrae, predominate, as well as for other Bering Sea and Northwest Pacific coast sites where salmon remains are frequently recovered (Butler and Campbell, 2004; S. K. Campbell and Butler, 2010; McKechnie and Moss, 2016).

Though the sample is unlikely to be fully representative of all salmon exploitation at Nunalleq, it nevertheless demonstrates that all species of Pacific salmon available within the study region were exploited. Applied to a larger sample from different types of deposits (i.e., house floors and storage pit fills), and occupation phases, rapid collagen-based species identification has the clear potential to elucidate salmon exploitation strategies, storage practices and food habits at the site, as well as

investigate potential changes in subsistence through time. The latter would be of particular interest in the context of the Little Ice Age (LIA), which likely affected salmon populations during severe climatic episodes (Finney et al., 2000). Zooarchaeological evidence has already confirmed that salmon exploitation likely decreased during the LIA at the site: whether or not this decrease was universal or related to the disappearance of certain species from the menu can now be a subject of future work.

Even from these initial results, however, the potential predominance of chum salmon (making up 30% of the salmonid assemblage) is intriguing as the species has a low fat content and was preferred ethnographically as a dried source of food stored in the winter for dogs and humans (Cannon and Yang, 2006; Seitz, 1990; Walker et al., 1993). Keeping dog teams required significant amounts of food, especially during the winter, and the amounts necessary are estimated to be similar to the requirements of humans (McManus-Fry et al., 2018; Morey and Aaris-Sørensen, 2002). The high proportion of chum identified could represent the feeding of dogs at the site, who were vital for traction and hunting (Masson-MacLean et al. in review). Future research could include the sampling of fish remains identified in dog coprolites (Masson-Maclean et al., n.d. In review) to test this hypothesis.

5. Conclusion

Through combined biomolecular methods including collagen sequencing, peptide mass fingerprinting, and ancient mtDNA analysis, we identify a series of collagen peptide markers that can distinguish Pacific salmon using MALDI-TOF, and validate these species-specific classifications through ancient mtDNA analysis. These collagen biomarkers provide a rapid and cost-effective method for distinguishing Pacific salmon within archaeological contexts, and further expand the applicability of ZooMS to archaeological fish bone assemblages from the Pacific Northwest. Considering the importance of marine resources in the lives of Indigenous cultures of the region, ZooMS has a high potential for increasing the taxonomic resolution for other morphologically cryptic fish taxa, including small forage fish like smelt, herring and eulachon (Palmer et al., 2018), and the (over 70) species of Northeast Pacific Rockfish (*Sebastes* spp.) (Rodrigues et al., 2018).

Acknowledgements

This research would not be possible without Qanirtuuq Incorporated (especially Warren Jones and Michael Smith), Quinhagak, Alaska, and the people of Quinhagak who

provided logistical and planning support during fieldwork and sampling permissions. Special thanks to all Nunalleq project team members, in Aberdeen and at other institutions, as well as volunteers, for their involvement in the excavations and their meticulous recovery of the faunal remains at the site. We are grateful to Dr Arturo Morales-Muñiz (Universidad Autónoma de Madrid), Dr Laura Llorente Rodriguez (Leiden University), Dr David Orton (University of York) and researchers at several conferences of the Fisheries Society of the British Isles for their insightful and productive conversations about salmon and to the two reviewers who commented on the manuscript. Thanks also to Dr. Laura Llorente Rodríguez for securing access to modern reference materials, and to Dr Madonna Moss (University of Oregon) for assistance in preparing the Kanektok River salmon and trout specimens for the University of Aberdeen reference collection. Thanks to Curtis J. Logan for assistance with some of the figures. This research was supported by funding through the Arts and Humanities Research Council (AH/K006029/1) grant awarded to KB, Rick Knecht, and Charlotta Hillerdal (University of Aberdeen), and a further AHRC-LabEx award (AH/N504543/1) to KB, KD, Rick Knecht (Aberdeen), and Isabelle Sidéra (Nanterre); by the Association for Environmental Archaeology Small Research Grants (EMM); BABAO (British Association for Biological Anthropology and Osteoarchaeology) Research Project Grant; the Royal Society Newton Fellowships (KKR); European Research Council Marie Skłodowska-Curie actions MAFRI (660337 to KKR); and the Leverhulme Trust (Philip Leverhulme Prize to CFS).

Data Statement

The MALDI spectra data produced through this study are available through Zenodo RN 2649336 (doi:10.5281/zenodo.2649336) while the LC-MS/MS data are available at the ProteomeXchange through MassIVE MSV000083687 (doi:10.25345/C5G04C). Ancient mtDNA sequences are available through the Genetic Sequence Database at the National Center for Biotechnical Information (NCBI) (GenBank ID: MK863994-MK864031).

References

Abadía-Cardoso, A., Garza, J.C., Mayden, R.L., García de León, F.J., 2015. Genetic Structure of Pacific Trout at the Extreme Southern End of Their Native Range. PLoS One 10, e0141775.

Alexandrou, M.A., Swartz, B.A., Matzke, N.J., Oakley, T.H., 2013. Genome duplication and multiple evolutionary origins of complex migratory behavior in Salmonidae. Mol. Phylogenet. Evol. 69, 514–523.

Allen, B.E., Anderson, M.L., Mee, J.A., Coombs, M., Rogers, S.M., 2016. Role of genetic

background in the introgressive hybridization of rainbow trout (Oncorhynchus mykiss) with Westslope cutthroat trout (O. clarkii lewisi). Conserv. Genet. 17, 521–531.
Allendorf, F.W., Thorgaard, G.H., 1984. Evolutionary genetics of fishes. Tetraploidy and the Evolution of Salmonid Fishes 55–93.
Barker, J.H., 1993. Always getting ready: Upterrlainarluta: Yup'ik Eskimo subsistence in southwest Alaska. University of Washington Press, Seattle.
Bern, M., Kil, Y.J., Becker, C., 2012. Byonic: Advanced Peptide and Protein Identification Software, in: Baxevanis, A.D., Petsko, G.A., Stein, L.D., Stormo, G.D. (Eds.), Current Protocols in Bioinformatics. John Wiley & Sons, Inc., Hoboken, NJ, USA, p. 1393.
Berthelot, C., Brunet, F., Chalopin, D., Juanchich, A., Bernard, M., Noël, B., Bento, P., Da Silva, C., Labadie, K., Alberti, A., Aury, J.-M., Louis, A., Dehais, P., Bardou, P., Montfort, J., Klopp, C., Cabau, C., Gaspin, C., Thorgaard, G.H., Boussaha, M., Quillet, E., Guyomard, R., Galiana, D., Bobe, J., Volff, J.-N., Genêt, C., Wincker, P., Jaillon, O., Roest Crollius, H., Guiguen, Y., 2014. The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. Nat. Commun. 5, 3657.
Biard, V., Gol'din, P., Gladilina, E., Vishnyakova, K., McGrath, K., Vieira, F.G., Wales, N., Fontaine, M.C., Speller, C., Olsen, M.T., 2017. Genomic and proteomic identification of Late Holocene remains: Setting baselines for Black Sea odontocetes. Journal of Archaeological Science: Reports 15, 262–271.
Britton, K., Knecht, R., Nehlich, O., Hillerdal, C., Davis, R.S., Richards, M.P., 2013. Maritime adaptations and dietary variation in prehistoric Western Alaska: Stable isotope analysis of permafrost-preserved human hair. Am. J. Phys. Anthropol. 151, 448–461.
Britton, K., McManus-Fry, E., Nehlich, O., Richards, M., Ledger, P.M., Knecht, R., 2018. Stable carbon, nitrogen and sulphur isotope analysis of permafrost preserved human hair from rescue excavations (2009, 2010) at the precontact site of Nunalleq, Alaska. Journal of Archaeological Science: Reports 17, 950–963.
Broughton, R.E., Betancur-R, R., Li, C., Arratia, G., Ortí, G., 2013. Multi-locus phylogenetic analysis reveals the pattern and tempo of bony fish evolution. PLoS Curr. 5. https://doi.org/10.1371/currents.tol.2ca8041495ffafd0c92756e75247483e
Brown, S., Higham, T., Slon, V., Pääbo, S., Meyer, M., Douka, K., Brock, F., Comeskey, D., Procopio, N., Shunkov, M., Derevianko, A., Buckley, M., 2016. Identification of a new hominin bone from Denisova Cave, Siberia using collagen fingerprinting and mitochondrial DNA analysis. Sci. Rep. 6, 23559.
Buckley, M., 2018. Zooarchaeology by Mass Spectrometry (ZooMS) Collagen Fingerprinting for the Species Identification of Archaeological Bone Fragments, in: Giovas, C.M., LeFebvre, M.J. (Eds.), Zooarchaeology in Practice: Case Studies in Methodology and Interpretation in Archaeofaunal Analysis. Springer International Publishing, Cham, pp. 227–247.
Buckley, M., Collins, M., Thomas-Oates, J., 2008. A method of isolating the collagen (I) alpha2 chain carboxytelopeptide for species identification in bone fragments. Anal. Biochem. 374, 325–334.
Buckley, M., Collins, M., Thomas-Oates, J., Wilson, J.C., 2009. Species identification by analysis of bone collagen using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom. 23, 3843–3854.
Buckley, M., Fraser, S., Herman, J., Melton, N.D., Mulville, J., Pálsdóttir, A.H., 2014. Species identification of archaeological marine mammals using collagen fingerprinting. J. Archaeol. Sci. 41, 631–641.
Buckley, M., Gu, M., Shameer, S., Patel, S., Chamberlain, A.T., 2016. High-throughput collagen fingerprinting of intact microfaunal remains; a low-cost method for distinguishing between

murine rodent bones. Rapid Commun. Mass Spectrom. 30, 805–812.
Butler, V.L., Bowers, N.J., 1998. Ancient DNA from salmon bone: a preliminary study. Anc. Biomol. 2, 17–26.
Butler, V.L., Campbell, S.K., 2004. Resource Intensification and Resource Depression in the Pacific Northwest of North America: A Zooarchaeological Review. J World Prehist 18, 327–405.
Campbell, L.A., 2010. Life histories of juvenile Chinook Salmon (Oncorhynchus tshawytscha) in the Columbia River estuary as inferred from scale and otolith microchemistry.
Campbell, M.A., López, J.A., Sado, T., Miya, M., 2013. Pike and salmon as sister taxa: detailed intraclade resolution and divergence time estimation of Esociformes + Salmoniformes based on whole mitochondrial genome sequences. Gene 530, 57–65.
Campbell, S., Butler, V., 2010. Archaeological Evidence for Resilience of Pacific Northwest Salmon Populations and the Socioecological System over the last ~7,500 years. Ecol. Soc. 15. https://doi.org/10.5751/ES-03151-150117
Campbell, S.K., Butler, V.L., 2010. Archaeological evidence for resilience of Pacific Northwest salmon populations and the socioecological system over the last\ 7,500 years. Ecol. Soc. 15.
Cannon, A., 1988. Radiographic Age Determination of Pacific Salmon: Species and Seasonal Inferences. J. Field Archaeol. 15, 103–108.
Cannon, A., Carlson, R.L., 1991. The economic prehistory of Namu: patterns in vertebrate fauna. Department of Archaeology and The Institute of Fisheries Analysis.
Cannon, A., Speller, C.F., Yang, D.Y., 2011. Site Specific Salmon Fisheries on the Central Coast of British Colombia, in: The Archaeology of North Pacific Fisheries. University of Alaska Press, pp. 57–74.
Cannon, A., Yang, D.Y., 2006. Early Storage and Sedentism on the Pacific Northwest Coast: Ancient DNA Analysis of Salmon Remains from Namu, British Columbia. Am. Antiq. 71, 123–140.
Casteel, R.W., 1974. Identification of the Species of Pacific Salmon (Genus Oncorhynchus) Native to North America Based upon Otoliths. Copeia 1974, 305–311.
Christoffels, A., Koh, E.G.L., Chia, J.-M., Brenner, S., Aparicio, S., Venkatesh, B., 2004. Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. Mol. Biol. Evol. 21, 1146–1151.
Collins, M., Buckley, M., Grundy, H.H., Thomas-Oates, J., Wilson, J., van Doorn, N., 2010. ZooMS: the collagen barcode and fingerprints. Spectrosc. Eur. 22, 2–2.
Crête-Lafrenière, A., Weir, L.K., Bernatchez, L., 2012. Framing the Salmonidae family phylogenetic portrait: a more complete picture from increased taxon sampling. PLoS One 7, e46662.
Dittman, A., Quinn, T., 1996. Homing in Pacific salmon: mechanisms and ecological basis. J. Exp. Biol. 199, 83–91.
Ewonus, P.A., Cannon, A., Yang, D.Y., 2011. Addressing seasonal site use through ancient DNA species identification of Pacific salmon at Dionisio Point, Galiano Island, British Columbia. J. Archaeol. Sci. 38, 2536–2546.
Farrell, T.F.G., Jordan, P., Taché, K., Lucquin, A., 2014. Specialized processing of aquatic resources in prehistoric Alaskan pottery? A lipid-residue analysis of ceramic sherds from the Thule-period site of Nunalleq, Alaska. Arctic.
Ferguson, M.M., Danzmann, R.G., Allendorf, F.W., 1985. Absence of developmental incompatibility in hybrids between rainbow trout and two subspecies of cutthroat trout. Biochem. Genet. 23, 557–570.

Fienup-Riordan, A., Rearden, A., 2016. Anguyiim Nalliini: The History of Bow-and-arrow Warfare in Southwest Alaska.
Fienup-Riordan, A., Rearden, A., Knecht, M., 2015. Irr'inarqellriit/Amazing things: Quinhagak elders reflect on their past. Alaska Journal of Anthropology 13, 37–71.
Finney, B.P., Gregory-Eaves, I., Sweetman, J., Douglas, M.S., Smol, J.P., 2000. Impacts of climatic change and fishing on Pacific salmon abundance over the past 300 years. Science 290, 795–799.
Frank, A., 2008. My Legacy to You. University of Washington Press, Seattle.
Friedland, K.D., Haas, R.E., 1996. Marine post-smolt growth and age at maturity of Atlantic salmon. J. Fish Biol. 48, 1–15.
Fukuwaka, M.-A., 1998. Scale and otolith patterns prove growth history of Pacific salmon. Bull. NPAFC 1, 190–198.
Funk, C., 2010. The bow and arrow war days on the Yukon-Kuskokwim delta of Alaska. Ethnohistory.
Garibaldi, A., Turner, N., 2004. Cultural Keystone Species: Implications for Ecological Conservation and Restoration. Ecology and Society 9, 1.
Gigleux, C., Grimes, V., Tütken, T., Knecht, R., Britton, K., 2019. Reconstructing caribou seasonal biogeography in Little Ice Age (late Holocene) Western Alaska using intra-tooth strontium and oxygen isotope analysis. Journal of Archaeological Science: Reports 23, 1043–1054.
Gistelinck, C., Gioia, R., Gagliardi, A., Tonelli, F., Marchese, L., Bianchi, L., Landi, C., Bini, L., Huysseune, A., Witten, P.E., Staes, A., Gevaert, K., De Rocker, N., Menten, B., Malfait, F., Leikin, S., Carra, S., Tenni, R., Rossi, A., De Paepe, A., Coucke, P., Willaert, A., Forlino, A., 2016. Zebrafish Collagen Type I: Molecular and Biochemical Characterization of the Major Structural Protein in Bone and Skin. Sci. Rep. 6, 21540.
Gobalet, K.W., 2001. A Critique of Faunal Analysis; Inconsistency among Experts in Blind Tests. J. Archaeol. Sci. 28, 377–386.
Gong, L., Liu, L.-Q., Guo, B.-Y., Ye, Y.-Y., Lü, Z.-M., 2017. The complete mitochondrial genome of Oncorhynchus masou formosanus (Salmoniformes: Salmonidae) and phylogenetic studies of Salmoninae. Conserv. Genet. Resour. 9, 281–284.
Grier, C., Flanigan, K., Winters, M., Jordan, L.G., Lukowski, S., Kemp, B.M., 2013. Using ancient DNA identification and osteometric measures of archaeological Pacific salmon vertebrae for reconstructing salmon fisheries and site seasonality at Dionisio Point, British Columbia. J. Archaeol. Sci. 40, 544–555.
Groot, G., 1991. Pacific Salmon Life Histories. UBC Press.
Gu, X., Li, W.H., 1992. Higher rates of amino acid substitution in rodents than in humans. Mol. Phylogenet. Evol. 1, 211–214.
Halffman, C.M., Potter, B.A., McKinney, H.J., Finney, B.P., Rodrigues, A.T., Yang, D.Y., Kemp, B.M., 2015. Early human use of anadromous salmon in North America at 11,500 y ago. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.1509747112
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.
Harris, W.A., Janecki, D.J., Reilly, J.P., 2002. Use of matrix clusters and trypsin autolysis fragments as mass calibrants in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom. 16, 1714–1722.
Harvey, V.L., Daugnora, L., Buckley, M., 2018. Species identification of ancient Lithuanian fish remains using collagen fingerprinting. J. Archaeol. Sci. 98, 102–111.
Hauser, B., 2014. Fishes of the Last Frontier: Life Histories, Biology, Ecology, and Management

of Alaska's Fishes. Publication Consultants.
Henry, K.A., 1961. Bulletin XII: Racial identification of Fraser River sockeye salmon by means of scales and its applications to salmon management. International Pacific Salmon Fisheries Commission.
Hickinbotham, S., Collins, M., 2017. Bacollite: analysis of MALDI and LC-MS/MS spectral data using sequence data (mainly for collagen). https://github.com/bioarch-sjh/bacollite.
Hickinbotham, S.J., Fiddyment, S., Stinson, T.L., Collins, M.J., n.d. How to Get Your Goat: Automated Identification of Species from MALDI-ToF Spectra. Bioinformatics Under Review.
Hillerdal, C., 2017. Integrating the past in the present: archaeology as part of living Yup'ik heritage, in: Archaeologies of "Us" and "Them." Routledge, pp. 82–97.
Hillerdal, C., Knecht, R., Jones, W., 2018. Nunalleq: Archaeology, Climate Change and Community Engagement in a Yup'ik Village. Arctic Anthropol. 56.
Hofkamp, A.R., Butler, V.L., 2017. On the validity of the radiographic method for determining age of ancient salmon. Journal of Archaeological Science: Reports 12, 449–456.
Hofman, C.A., Rick, T.C., Erlandson, J.M., Reeder-Myers, L., Welch, A.J., Buckley, M., 2018. Collagen Fingerprinting and the Earliest Marine Mammal Hunting in North America. Sci. Rep. 8, 10014.
Horreo, J.L., 2017. Revisiting the mitogenomic phylogeny of Salmoninae: new insights thanks to recent sequencing advances. PeerJ 5, e3828.
Huber, H.R., Jorgensen, J.C., Butler, V.L., Baker, G., Stevens, R., 2011. Can salmonids (Oncorhynchus spp.) be identified to species using vertebral morphometrics? J. Archaeol. Sci. 38, 136–146.
Hurley, I.A., Mueller, R.L., Dunn, K.A., Schmidt, E.J., Friedman, M., Ho, R.K., Prince, V.E., Yang, Z., Thomas, M.G., Coates, M.I., 2006. A new time-scale for ray-finned fish evolution. Proceedings of the Royal Society B: Biological Sciences 274, 489–498.
Keller, B.O., Sui, J., Young, A.B., Whittal, R.M., 2008. Interferences and contaminants encountered in modern mass spectrometry. Anal. Chim. Acta 627, 71–81.
Kemp, B.M., Monroe, C., Judd, K.G., Reams, E., Grier, C., 2014. Evaluation of methods that subdue the effects of polymerase chain reaction inhibitors in the study of ancient and degraded DNA. J. Archaeol. Sci. 42, 373–380.
Kimura, S., Miyauchi, Y., Uchida, N., 1991. Scale and bone type I collagens of carp (Cyprinus carpio). Comp. Biochem. Physiol. B 99, 473–476.
Kimura, S., Ohno, Y., 1987. Fish type I collagen: tissue-specific existence of two molecular forms, (α1)2α2 and α1α2α3, in Alaska pollack. Comparative Biochemistry and Physiology Part B: Comparative Biochemistry 88, 409–413.
Kitano, T., Matsuoka, N., Saitou, N., 1997. Phylogenetic relationship of the genus Oncorhynchus species inferred from nuclear and mitochondrial markers. Genes Genet. Syst. 72, 25–34.
Korzow-Richter, K.K., Wilson, J., Jones, A.K.G., Buckley, M., van Doorn, N., Collins, M.J., 2011. Fish 'n chips: ZooMS peptide mass fingerprinting in a 96 well plate format to identify fish bone fragments. J. Archaeol. Sci. 38, 1502–1510.
Kovach, R.P., Al-Chokhachy, R., Stephens, T., 2018. Proactive Rainbow Trout Suppression Reduces Threat of Hybridization in the Upper Snake River Basin. N. Am. J. Fish. Manage. 38, 811–819.
LaVine, R., Lisac, M.J., Coiley-Kenner, P., 2007. Traditional Ecological Knowledge of 20th-Century Ecosystems and Fish Populations in the Kuskokwim Bay Region. US Fish and Wildlife Service, Fisheries Resource Monitoring Program.

Lecaudey, L.A., Schliewen, U.K., Osinov, A.G., Taylor, E.B., Bernatchez, L., Weiss, S.J., 2018. Inferring phylogenetic structure, hybridization and divergence times within Salmoninae (Teleostei: Salmonidae) using RAD-sequencing. Mol. Phylogenet. Evol. 124, 82–99.

Ledger, P.M., Forbes, V., Masson-Maclean, E., Hillerdal, C., Derek Hamilton, W., McManus-Fry, E., Jorge, A., Britton, K., Knecht, R.A., 2018. Three generations under one roof? Bayesian modeling of radiocarbon data from Nunalleq, Yukon-Kuskokwim Delta, Alaska. Am. Antiq. 83, 505–524.

Ledger, P.M., Forbes, V., Masson-MacLean, E., Knecht, R.A., 2016. Dating and digging stratified archaeology in circumpolar North America: A view from Nunalleq, Southwestern Alaska. Arctic.

LeFebvre, M.J., Sharpe, A.E., 2018. Contemporary Challenges in Zooarchaeological Specimen Identification, in: Giovas, C.M., LeFebvre, M.J. (Eds.), Zooarchaeology in Practice: Case Studies in Methodology and Interpretation in Archaeofaunal Analysis. Springer International Publishing, Cham, pp. 35–57.

Martin, A.P., Palumbi, S.R., 1993. Body size, metabolic rate, generation time, and the molecular clock. Proc. Natl. Acad. Sci. U. S. A. 90, 4087–4091.

Masson-MacLean, E., Houmard, C., Knecht, R., Sidéra, I., Dobney, K., Britton, K., 2019. Pre-contact Adaptations to the Little Ice Age in Southwestern Alaska: New Zooarchaeological Evidence from the Nunalleq Site (15th–17th C. AD). Quat. Int.

Masson-Maclean, E., Mcmanus-Fry, E., Britton, K., n.d. The archaeology of dogs at the precontact Yup'ik site of Nunalleq, Western Alaska, in: Burtt, A., Bethke, D. (Eds.), Beyond Domestication: Archaeological Investigations into the Human-Canine Connection. University of Florida Press, Florida.

McKechnie, I., Moss, M.L., 2016. Meta-analysis in zooarchaeology expands perspectives on Indigenous fisheries of the Northwest Coast of North America. Journal of Archaeological Science: Reports 8, 470–485.

McManus-Fry, E., Knecht, R., Dobney, K., Richards, M.P., Britton, K., 2018. Dog-human dietary relationships in Yup'ik western Alaska: The stable isotope and zooarchaeological evidence from pre-contact Nunalleq. Journal of Archaeological Science: Reports 17, 964–972.

McMurrich, J.P., 1913. The Life Cycles of the Pacific Coast Salmon Belonging to the Genus Oncorhynchus: As Revealed by Their Scale and Otolith Markings. Royal Society of Canada.

Meyer, A., Van de Peer, Y., 2005. From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). Bioessays 27, 937–945.

Monroe, C., Grier, C., Kemp, B.M., 2013. Evaluating the efficacy of various thermo-stable polymerases against co-extracted PCR inhibitors in ancient DNA samples. Forensic Sci. Int. 228, 142–153.

Morey, D.F., Aaris-Sørensen, K., 2002. Paleoeskimo dogs of the eastern Arctic. Arctic 44–56.

Morvan-Dubois, G., Le Guellec, D., Garrone, R., Zylberberg, L., Bonnaud, L., 2003. Phylogenetic Analysis of Vertebrate Fibrillar Collagen Locates the Position of Zebrafish α3(I) and Suggests an Evolutionary Link Between Collagen α Chains and Hox Clusters. J. Mol. Evol. 57, 501–514.

Moss, M.L., Judd, K.G., Kemp, B.M., 2014. Can salmonids (Oncorhynchus spp.) be identified to species using vertebral morphometrics? A test using ancient DNA from Coffman Cove, Alaska. J. Archaeol. Sci. 41, 879–889.

Moss, M.L., Rodrigues, A.T., Speller, C.F., Yang, D.Y., 2016/8. The historical ecology of Pacific herring: Tracing Alaska Native use of a forage fish. Journal of Archaeological Science: Reports 8, 504–512.

Murray, C.B., 1994. A Method for Preparing Chinook Salmon Otoliths for Age Determination, and Evidence of Its Validity. Trans. Am. Fish. Soc. 123, 358–367.
Near, T.J., Eytan, R.I., Dornburg, A., Kuhn, K.L., Moore, J.A., Davis, M.P., Wainwright, P.C., Friedman, M., Smith, W.L., 2012. Resolution of ray-finned fish phylogeny and timing of diversification. Proc. Natl. Acad. Sci. U. S. A. 109, 13698–13703.
Palmer, E., Tushingham, S., Kemp, B.M., 2018. Human use of small forage fish: Improved ancient DNA species identification techniques reveal long term record of sustainable mass harvesting of smelt fishery in the northeast Pacific Rim. J. Archaeol. Sci. 99, 143–152.
Pasquier, J., Cabau, C., Nguyen, T., Jouanno, E., Severac, D., Braasch, I., Journot, L., Pontarotti, P., Klopp, C., Postlethwait, J.H., Guiguen, Y., Bobe, J., 2016. Gene evolution and gene expression after whole genome duplication in fish: the PhyloFish database. BMC Genomics 17, 368.
Penaluna, B.E., Abadía-Cardoso, A., Dunham, J.B., García-Dé León, F.J., Gresswell, R.E., Luna, A.R., Taylor, E.B., Shepard, B.B., Al-Chokhachy, R., Muhlfeld, C.C., Bestgen, K.R., Rogers, K., Escalante, M.A., Keeley, E.R., Temple, G.M., Williams, J.E., Matthews, K.R., Pierce, R., Mayden, R.L., Kovach, R.P., Garza, J.C., Fausch, K.D., 2016. Conservation of Native Pacific Trout Diversity in Western North America. Fisheries 41, 286–300.
Petre, S.J., Bonar, S.A., 2017. Determination of Habitat Requirements for Apache Trout. Trans. Am. Fish. Soc. 146, 1–15.
Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.
Prendergast, M.E., Buckley, M., Crowther, A., Frantz, L., Eager, H., Lebrasseur, O., Hutterer, R., Hulme-Beaman, A., Van Neer, W., Douka, K., Veall, M.-A., Quintana Morales, E.M., Schuenemann, V.J., Reiter, E., Allen, R., Dimopoulos, E.A., Helm, R.M., Shipton, C., Mwebi, O., Denys, C., Horton, M., Wynne-Jones, S., Fleisher, J., Radimilahy, C., Wright, H., Searle, J.B., Krause, J., Larson, G., Boivin, N.L., 2017. Reconstructing Asian faunal introductions to eastern Africa from multi-proxy biomolecular and archaeological datasets. PLoS One 12, e0182565.
Quinn, 2018. The Behavior and Ecology of Pacific Salmon and Trout, 2 edition. ed. University of Washington Press.
Quinn, T.P., Dittman, A.H., 1990. Pacific salmon migrations and homing: mechanisms and adaptive significance. Trends Ecol. Evol. 5, 174–177.
Raghavan, M., DeGiorgio, M., Albrechtsen, A., Moltke, I., Skoglund, P., Korneliussen, T.S., Grønnow, B., Appelt, M., Gulløv, H.C., Friesen, T.M., Fitzhugh, W., Malmström, H., Rasmussen, S., Olsen, J., Melchior, L., Fuller, B.T., Fahrni, S.M., Stafford, T., Jr, Grimes, V., Renouf, M.A.P., Cybulski, J., Lynnerup, N., Lahr, M.M., Britton, K., Knecht, R., Arneborg, J., Metspalu, M., Cornejo, O.E., Malaspinas, A.-S., Wang, Y., Rasmussen, M., Raghavan, V., Hansen, T.V.O., Khusnutdinova, E., Pierre, T., Dneprovsky, K., Andreasen, C., Lange, H., Hayes, M.G., Coltrain, J., Spitsyn, V.A., Götherström, A., Orlando, L., Kivisild, T., Villems, R., Crawford, M.H., Nielsen, F.C., Dissing, J., Heinemeier, J., Meldgaard, M., Bustamante, C., O'Rourke, D.H., Jakobsson, M., Gilbert, M.T.P., Nielsen, R., Willerslev, E., 2014. The genetic prehistory of the New World Arctic. Science 345, 1255832.
Rambaut, A., 2007. FigTree, a graphical viewer of phylogenetic trees. See http://tree. bio. ed. ac. uk/software/figtree.
Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. Syst. Biol. 67, 901–904.
Rasmussen, R.S., Morrissey, M.T., Hebert, P.D.N., 2009. DNA barcoding of commercially

important salmon and trout species (Oncorhynchus and Salmo) from North America. J. Agric. Food Chem. 57, 8379–8385.
Rodrigues, A.T., McKechnie, I., Yang, D.Y., 2018. Ancient DNA analysis of Indigenous rockfish use on the Pacific Coast: Implications for marine conservation areas and fisheries management. PLoS One 13, e0192716.
Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Rosenfield, J.A., Todd, T., Greil, R., 2000. Asymmetric Hybridization and Introgression between Pink Salmon and Chinook Salmon in the Laurentian Great Lakes. Trans. Am. Fish. Soc. 129, 670–679.
Royle, T.C.A., Sakhrani, D., Speller, C.F., Butler, V.L., Devlin, R.H., Cannon, A., Yang, D.Y., 2018. An efficient and reliable DNA-based sex identification method for archaeological Pacific salmonid (Oncorhynchus spp.) remains. PLoS One 13, e0193212.
Saglam, I.K., Prince, D.J., Meek, M., Ali, O.A., Miller, M.R., Peacock, M., Neville, H., Goodbla, A., Mellison, C., Somer, W., May, B., Finger, A.J., 2017. Genomic Analysis Reveals Genetic Distinctiveness of the Paiute Cutthroat Trout Oncorhynchus clarkii seleniris. Trans. Am. Fish. Soc. 146, 1291–1302.
Saito, T., Kaga, T., Seki, J., Otake, T., 2007. Otolith microstructure of chum salmon Oncorhynchus keta: formation of sea entry check and daily deposition of otolith increments in seawater conditions. Fish. Sci. 73, 27.
Santini, F., Harmon, L.J., Carnevale, G., Alfaro, M.E., 2009. Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. BMC Evol. Biol. 9, 194.
Seitz, J., 1990. Subsistence Salmon Fishing in Nushagak Bay, Southwest Alaska. Alaska Department of Fish and Game, Division of Subsistence.
Song, L., Wangs, W., 2013. Genomes and evolutionary genomics of animals. Curr. Zool. 59, 87–98.
Speller, C.F., Hauser, L., Lepofsky, D., Moore, J., Rodrigues, A.T., Moss, M.L., McKechnie, I., Yang, D.Y., 2012. High potential for using DNA from ancient herring bones to inform modern fisheries management and conservation. PLoS One 7, e51122.
Speller, C.F., Yang, D.Y., Hayden, B., 2005. Ancient DNA investigation of prehistoric salmon resource utilization at Keatley Creek, British Columbia, Canada. J. Archaeol. Sci. 32, 1378–1389.
Stapels, M.D., Barofsky, D.F., 2004. Complementary use of MALDI and ESI for the HPLC-MS/MS analysis of DNA-binding proteins. Anal. Chem. 76, 5423–5430.
Star, B., Boessenkool, S., Gondek, A.T., Nikulina, E.A., Hufthammer, A.K., Pampoulie, C., Knutsen, H., André, C., Nistelberger, H.M., Dierking, J., Petereit, C., Heinrich, D., Jakobsen, K.S., Stenseth, N.C., Jentoft, S., Barrett, J.H., 2017. Ancient DNA reveals the Arctic origin of Viking Age cod from Haithabu, Germany. Proc. Natl. Acad. Sci. U. S. A. 114, 9152–9157.
Strohalm, M., Hassman, M., Kosata, B., Kodícek, M., 2008. mMass data miner: an open source alternative for mass spectrometric data analysis. Rapid Commun. Mass Spectrom. 22, 905–908.
Takezaki, N., 2018. Global Rate Variation in Bony Vertebrates. Genome Biol. Evol. 10, 1803–1815.
Taylor, W., Shnaider, S., Abdykanova, A., Fages, A., Welker, F., Irmer, F., Seguin-Orlando, A., Khan, N., Douka, K., Kolobova, K., Orlando, L., Krivoshapkin, A., Boivin, N., 2018. Early pastoral economies along the Ancient Silk Road: Biomolecular evidence from the Alay

Valley, Kyrgyzstan. PLoS One 13, e0205646.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.
Thompson, T.Q., Bellinger, M.R., O'Rourke, S.M., Prince, D.J., Stevenson, A.E., Rodrigues, A.T., Sloat, M.R., Speller, C.F., Yang, D.Y., Butler, V.L., Banks, M.A., Miller, M.R., 2019. Anthropogenic habitat alteration leads to rapid loss of adaptive variation and restoration potential in wild salmon populations. Proc. Natl. Acad. Sci. U. S. A. 116, 177–186.
von Holstein, I.C.C., Ashby, S.P., van Doorn, N.L., Sachs, S.M., Buckley, M., Meiri, M., Barnes, I., Brundle, A., Collins, M.J., 2014. Searching for Scandinavians in pre-Viking Scotland: molecular fingerprinting of Early Medieval combs. J. Archaeol. Sci. 41, 1–6.
Walker, R.J., Coffing, M., Alaska, J., 1993. Subsistence salmon harvests in the Kuskokwim area during 1989. Alaska Department of Fish and Game, Division of Subsistence.
Welker, F., Collins, M.J., Thomas, J.A., Wadsley, M., Brace, S., Cappellini, E., Turvey, S.T., Reguero, M., Gelfo, J.N., Kramarz, A., Burger, J., Thomas-Oates, J., Ashford, D.A., Ashton, P.D., Rowsell, K., Porter, D.M., Kessler, B., Fischer, R., Baessmann, C., Kaspar, S., Olsen, J.V., Kiley, P., Elliott, J.A., Kelstrup, C.D., Mullin, V., Hofreiter, M., Willerslev, E., Hublin, J.-J., Orlando, L., Barnes, I., MacPhee, R.D.E., 2015a. Ancient proteins resolve the evolutionary history of Darwin's South American ungulates. Nature 522, 81–84.
Welker, F., Soressi, M., Rendu, W., Hublin, J.-J., Collins, M., 2015b. Using ZooMS to identify fragmentary bone from the Late Middle/Early Upper Palaeolithic sequence of Les Cottés, France. J. Archaeol. Sci. 54, 279–286.
Willson, M.F., Halupka, K.C., 1995. Anadromous Fish as Keystone Species in Vertebrate Communities. Conserv. Biol. 9, 489–497.
Wolfe, R.J., Gross, J.J., Langdon, G.J., Wright, J.M., Sherrod, G.K., Ellanna, L.J., Sumida, V., Usher, P.J., 1984. Subsistence-based economies in coastal communities of southwest Alaska, Technical Paper No. 89. Alaska Department of Fish and Game, Subsistence Division. Anchorage, Alaska.
Yang, D.Y., Eng, B., Waye, J.S., Dudar, J.C., Saunders, S.R., 1998. Technical note: improved DNA extraction from ancient bones using silica-based spin columns. Am. J. Phys. Anthropol. 105, 539–543.
Yang, D.Y., Speller, C.F., 2006. Co-amplification of cytochrome b and D-loop mtDNA fragments for the identification of degraded DNA samples. Mol. Ecol. Notes 6, 605–608.
Zhivotovsky, L.A., 2015. Genetic history of salmonid fishes of the genus Oncorhynchus. Russ. J. Genet. 51, 491–505.